



Title	Kernel-Induced Sampling Theorem
Author(s)	Tanaka, Akira; Imai, Hideyuki; Miyakoshi, Masaaki
Citation	IEEE Transactions on Signal Processing, 58(7), 3569-3577 <a href="https://doi.org/10.1109/TSP.2010.2046637">https://doi.org/10.1109/TSP.2010.2046637</a>
Issue Date	2010-07
Doc URL	<a href="http://hdl.handle.net/2115/43185">http://hdl.handle.net/2115/43185</a>
Rights	© 2010 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.
Type	article
File Information	TSP58-7_3569-3577.pdf



[Instructions for use](#)

# Kernel-Induced Sampling Theorem

Akira Tanaka, Hideyuki Imai, and Masaaki Miyakoshi

**Abstract**—A perfect reconstruction of functions in a reproducing kernel Hilbert space from a given set of sampling points is discussed. A necessary and sufficient condition for the corresponding reproducing kernel and the given set of sampling points to perfectly recover the functions is obtained in this paper. The key idea of our work is adopting the reproducing kernel Hilbert space corresponding to the Gramian matrix of the kernel and the given set of sampling points as the range space of a sampling operator and considering the orthogonal projector, defined via the range space, onto the closed linear subspace spanned by the kernel functions corresponding to the given sampling points. We also give an error analysis of a reconstructed function by incomplete sampling points.

**Index Terms**—Gramian matrix, Hilbert space, orthogonal projection, reproducing kernel, sampling theorem.

## I. INTRODUCTION

SHANNON's sampling theorem [1] claims that

$$f(x) = \sum_{n=-\infty}^{\infty} f(n) \frac{\sin \pi(x-n)}{\pi(x-n)} \quad (1)$$

holds for any  $\pi$ -bandlimited function  $f(x)$  with finite energy. This theorem plays a crucial role not only in the field of signal processing but also in many other scientific areas. There exist so many generalizations and extensions of this theorem such as nonuniform sampling (see [2], [3], and their references cited in). Among them, a reformulation of sampling theorem, using a reproducing kernel Hilbert space (RKHS), presented by Nashed and Walter [4] is one of the most important milestones in the history of the sampling theorem. Their formulation enables us to obtain a rigorous description of a sampling process by using the reproducing property of a reproducing kernel and gives us a unified viewpoint for many generalizations and extensions of the sampling theorem including the subsequent wavelet-based sampling theories [2], [5]. In the literature, they discussed the relationship between a given system of kernel functions corresponding to a given set of sampling points and a reproducing kernel (or RKHS), which leads a perfect reconstruction of any function in the RKHS (or its subspaces). However, we do not have a necessary and sufficient condition, that is easy to check

in a general and practical case, for a given system of kernel functions corresponding to a given set of sampling points to be complete in the corresponding RKHS.

On the other hand, as one of different approaches of a generalization of the sampling theorem using a RKHS, Ogawa [6], [7] and Hirabayashi *et al.* [8] introduced a framework of the optimal approximation of a function in the RKHS, instead of a perfect reconstruction. The key idea of this framework is the orthogonal projection onto the linear subspace spanned by the given system of kernel functions corresponding to sampling points which may be incomplete for the RKHS. However, this framework is only for a finite number of sampling points. It is caused by an ad hoc treatment of the range space of a sampling operator.

In this paper, we extend the framework of [6]–[8] to infinite sampling points by adopting the RKHS corresponding to the Gramian matrix of the given kernel and a given set of infinite sampling points as the range space of a sampling operator; and on the basis of the extension, we give a necessary and sufficient condition for the kernel and the given set of sampling points to obtain the sampling theorem for the RKHS corresponding to the adopted kernel. We also give an error analysis for incomplete sampling points. Moreover, on the basis of our results, we show another proof of Shannon's sampling theorem and introduce a sampling theorem for a RKHS corresponding to a polynomial kernel; we also show that Sobolev spaces do not have a sampling theorem by equally spaced sampling points; and also show a numerical example of the reconstruction error with incomplete sampling points for the RKHS corresponding to the Gaussian kernel.

## II. MATHEMATICAL PRELIMINARIES FOR THE THEORY OF REPRODUCING KERNEL HILBERT SPACES

In this section, we prepare some mathematical tools concerned with the theory of reproducing kernel Hilbert spaces [9]–[11].

**Definition 1:** [9] Let  $\mathbf{R}^n$  be an  $n$ -dimensional real vector space and let  $\mathcal{H}$  be a class of functions defined on  $D \subset \mathbf{R}^n$ , forming a Hilbert space of real-valued functions. The function  $K(\mathbf{x}, \tilde{\mathbf{x}})$ , ( $\mathbf{x}, \tilde{\mathbf{x}} \in D$ ) is called a reproducing kernel of  $\mathcal{H}$ , if

1) For every fixed  $\tilde{\mathbf{x}} \in D$

$$K_{\tilde{\mathbf{x}}}(\cdot) = K(\cdot, \tilde{\mathbf{x}}) \in \mathcal{H}. \quad (2)$$

2) For every  $\tilde{\mathbf{x}} \in D$  and every  $f(\cdot) \in \mathcal{H}$ ,

$$f(\tilde{\mathbf{x}}) = \langle f(\cdot), K(\cdot, \tilde{\mathbf{x}}) \rangle_{\mathcal{H}} \quad (3)$$

where  $\langle \cdot, \cdot \rangle_{\mathcal{H}}$  denotes the inner product of the Hilbert space  $\mathcal{H}$ .

The Hilbert space  $\mathcal{H}$  that has a reproducing kernel  $K$  is called a reproducing kernel Hilbert space (RKHS), denoted by  $\mathcal{H}_K$ . The reproducing property (3) enables us to treat a value of a function at a point in  $D$ , while we can not deal with a value

Manuscript received November 29, 2009; accepted March 04, 2010. Date of publication March 25, 2010; date of current version June 16, 2010. This work was supported in part by a Grant-in-Aid No. 21700001 for Young Scientist (B) of the Ministry of Education, Culture, Sports and Technology of Japan. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Haldun M. Ozaktas.

The authors are with the Division of Computer Science, Graduate School of Information Science and Technology, Hokkaido University, Sapporo, 060-0814 Japan (e-mail: takira@main.eng.hokudai.ac.jp; imai@main.ist.hokudai.ac.jp; miyakosi@main.ist.hokudai.ac.jp).

Digital Object Identifier 10.1109/TSP.2010.2046637

of a function in a general Hilbert space such as  $L^2$ . Note that reproducing kernels are positive definite [9]

$$\sum_{i,j=1}^N c_i c_j K(\mathbf{x}_i, \mathbf{x}_j) \geq 0 \quad (4)$$

for any  $N, c_1, \dots, c_N \in \mathbf{R}$ , and  $\mathbf{x}_1, \dots, \mathbf{x}_N \in D$ . In addition,  $K(\mathbf{x}, \tilde{\mathbf{x}}) = K(\tilde{\mathbf{x}}, \mathbf{x})$  for any  $\mathbf{x}, \tilde{\mathbf{x}} \in D$  is followed [9]. If a reproducing kernel  $K(\mathbf{x}, \tilde{\mathbf{x}})$  exists, it is unique [9]. Conversely, every positive definite function  $K(\mathbf{x}, \tilde{\mathbf{x}})$  has the unique corresponding RKHS [9]. The following lemma is one of important properties of a RKHS.

*Lemma 1:*  $\{K(\cdot, \mathbf{x}) | \mathbf{x} \in D\}$  is complete in  $\mathcal{H}_K$ .

*Proof:* Let  $f(\cdot)$  be an arbitrary function in  $\mathcal{H}_K$ , which is orthogonal to  $K(\cdot, \mathbf{x})$  for any  $\mathbf{x} \in D$ . Then the reproducing property (3) and the orthogonality yield

$$f(\mathbf{x}) = \langle f(\cdot), K(\cdot, \mathbf{x}) \rangle_{\mathcal{H}_K} = 0$$

for any  $\mathbf{x} \in D$ , which concludes the proof. ■

Next, we introduce the Schatten product [12] that is a convenient tool to reveal the reproducing property of kernels.

*Definition 2:* [12] Let  $\mathcal{H}_1$  and  $\mathcal{H}_2$  be Hilbert spaces. The Schatten product of  $g \in \mathcal{H}_2$  and  $h \in \mathcal{H}_1$  is defined by

$$(g \otimes h)f = \langle f, h \rangle_{\mathcal{H}_1} g, \quad f \in \mathcal{H}_1. \quad (5)$$

Note that  $(g \otimes h)$  is a linear operator from  $\mathcal{H}_1$  onto  $\mathcal{H}_2$ . It is easy to show that the following relations hold for  $h, v \in \mathcal{H}_1$ ,  $g, u \in \mathcal{H}_2$ :

$$(h \otimes g)^* = (g \otimes h) \quad (6)$$

$$(h \otimes g)(u \otimes v) = \langle u, g \rangle_{\mathcal{H}_2} (h \otimes v) \quad (7)$$

where the superscript  $*$  denotes the adjoint operator.

### III. RKHS-BASED FORMULATION OF SAMPLING PROCESS AND OPTIMAL APPROXIMATION BY ORTHOGONAL PROJECTION

In this section, we formulate the sampling process of a function by using a reproducing kernel Hilbert space and discuss the orthogonal projection of the function onto the closed linear subspace spanned by the basis functions corresponding to sampling points. These discussions are basically along with the framework of [6]–[8] with an extension to infinite sampling points.

Let  $f(\cdot)$  be an arbitrary real-valued function belonging to some class of functions defined on  $D \subset \mathbf{R}^n$  and let  $X = \{\mathbf{x}_i \in D | i \in \mathbf{N}\}$  be a set of sampling points, where  $\mathbf{N}$  denotes the set of natural numbers.

The goal of sampling theorem is to clarify a necessary and sufficient condition to perfectly reconstruct the function  $f(\cdot)$  by using the function values at each point in  $X$  and some basis functions specified by  $X$ .

In this paper, we concentrate on the RKHS  $\mathcal{H}_K$  corresponding to some reproducing kernel  $K$  as a class of function to which the target functions belong. According to the reproducing property (3)

$$f(\mathbf{x}_i) = \langle f(\cdot), K(\cdot, \mathbf{x}_i) \rangle_{\mathcal{H}_K} \quad (8)$$

is obtained. Let  $\mathbf{e}_i$  be the unit vector in  $\mathbf{R}^\infty$  with only the  $i$ th component being unity and let  $\mathbf{f} = [f(\mathbf{x}_1), \dots, f(\mathbf{x}_\ell), \dots]' \in \mathbf{R}^\infty$  with  $'$  denoting the transposition operator. Then, (8) is rewritten as

$$\mathbf{f} = \left( \sum_{i=1}^{\infty} [\mathbf{e}_i \otimes K(\cdot, \mathbf{x}_i)] \right) f(\cdot) \quad (9)$$

by using the Schatten product [6]–[8]. For a convenience of description, we write

$$A = \left( \sum_{i=1}^{\infty} [\mathbf{e}_i \otimes K(\cdot, \mathbf{x}_i)] \right). \quad (10)$$

$A$  is a linear operator defined by  $K$  and  $X$  that maps an element of  $\mathcal{H}_K$  onto  $\mathbf{R}^\infty$ ; and (9) can be rewritten as

$$\mathbf{f} = A f(\cdot) \quad (11)$$

which represents the sampling process of  $f(\cdot) \in \mathcal{H}_K$  with the sampling points  $X$ . Therefore, function reconstruction process can be regarded as an inversion problem for (11) [6]–[8].

Here, we introduce three other spaces defined as follows.

Let  $S$  be a closed linear subspace in  $\mathcal{H}_K$ , spanned by the basis functions  $\{K(\cdot, \mathbf{x}_i) | i \in \mathbf{N}\}$ , defined as

$$S = \overline{\text{span}[\{K(\cdot, \mathbf{x}_i) | i \in \mathbf{N}\}]}. \quad (12)$$

Note that  $S^\perp = \mathcal{N}(A)$  holds, where  $S^\perp$  and  $\mathcal{N}(A)$  denote the orthogonal complement of  $S$  in  $\mathcal{H}_K$  and the null space of  $A$ , respectively. Any function  $f(\cdot) \in S$  can be represented by

$$f(\cdot) = \sum_{i=1}^{\infty} \alpha_i K(\cdot, \mathbf{x}_i) \quad (13)$$

with coefficients  $\alpha_i \in \mathbf{R}$ . Let  $\|\cdot\|_{\mathcal{H}_K}$  denotes the induced norm in  $\mathcal{H}_K$ , then, for any  $f(\cdot) \in S$

$$\begin{aligned} \|f(\cdot)\|_{\mathcal{H}_K}^2 &= \langle f(\cdot), f(\cdot) \rangle_{\mathcal{H}_K} \\ &= \left\langle \sum_{i=1}^{\infty} \alpha_i K(\cdot, \mathbf{x}_i), \sum_{j=1}^{\infty} \alpha_j K(\cdot, \mathbf{x}_j) \right\rangle_{\mathcal{H}_K} \\ &= \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \alpha_i \alpha_j \langle K(\cdot, \mathbf{x}_i), K(\cdot, \mathbf{x}_j) \rangle_{\mathcal{H}_K} \\ &= \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \alpha_i \alpha_j K(\mathbf{x}_i, \mathbf{x}_j) \\ &= \boldsymbol{\alpha}' G \boldsymbol{\alpha} < \infty \end{aligned}$$

holds, where  $\boldsymbol{\alpha} = [\alpha_1, \dots, \alpha_\ell, \dots]' \in \mathbf{R}^\infty$  and  $G = (K(\mathbf{x}_i, \mathbf{x}_j)) \in \mathbf{R}^{\infty \times \infty}$  denotes the Gramian matrix of the kernel  $K$  with sampling points  $X$ . We intend to use  $S$  as a linear subspace to which a reconstructed function belongs. Note that since  $S$  is closed

$$\mathcal{H}_C = \overline{\{\boldsymbol{\alpha} \in \mathbf{R}^\infty | \boldsymbol{\alpha}' G \boldsymbol{\alpha} < \infty\}} \quad (14)$$

is also a Hilbert space which is homeomorphic with  $S$ .

Here, we introduce the theorem shown by Aronszajn for the properties of  $G$ .

**Theorem 1:** [9] If  $K$  is the reproducing kernel of the class  $F$  of functions defined in the set  $E$  with the norm  $\|\cdot\|$ , then  $K$  restricted to a subset  $E_1 \subset E$  is the reproducing kernel of the class  $F_1$  of all restrictions of functions of  $F$  to the subset  $E_1$ . For any such restriction  $f_1 \in F_1$ , the norm  $\|f_1\|_1$  is the minimum of  $\|f\|$  for all  $f \in F$  whose restriction to  $E_1$  is  $f_1$ .

According to Theorem 1, it is concluded that  $G$  is also a reproducing kernel since  $X \subset D$ . Thus,  $G$  has the unique corresponding RKHS denoted by  $\mathcal{H}_G$ . Note that  $\mathcal{H}_G$  is a subset of  $\mathbf{R}^\infty$ . Also note that since  $\mathcal{H}_G$  is a Hilbert space, it is complete and closed, which implies that there exist a symmetric and non-negative matrix  $M$  that specifies the metric of  $\mathcal{H}_G$ . Thus,  $\mathcal{H}_G$  is characterized as

$$\mathcal{H}_G = \{\mathbf{g} \in \mathbf{R}^\infty | \mathbf{g}' M \mathbf{g} < \infty\}. \quad (15)$$

According to (2) in Definition 1

$$G\mathbf{e}_i \in \mathcal{H}_G \quad (16)$$

holds for any  $i \in \mathbf{N}$ , which implies that each column of  $G$  belongs to  $\mathcal{H}_G$ ; and (3) in Definition 1 yields

$$g_k = \langle \mathbf{g}, G\mathbf{e}_k \rangle_{\mathcal{H}_G} = \mathbf{e}_k' G M \mathbf{g} \quad (17)$$

for any  $\mathbf{g} = [g_1, \dots, g_\ell, \dots]' \in \mathcal{H}_G$ . The summation of (17), premultiplied by  $\mathbf{e}_k$ , with respect to  $k$  produces

$$\mathbf{g} = G M \mathbf{g} \quad (18)$$

for any  $\mathbf{g} \in \mathcal{H}_G$ . Therefore, since  $G\mathbf{e}_i \in \mathcal{H}_G$

$$G\mathbf{e}_i = G M G\mathbf{e}_i \quad (19)$$

holds for any  $i \in \mathbf{N}$ ; and the summation of (19), postmultiplied by  $\mathbf{e}_i'$ , with respect to  $i$  yields

$$G = G M G \quad (20)$$

which implies that  $M$  is a 1-inverse [13] of  $G$ .<sup>1</sup> Note that (20) for  $\mathcal{H}_G$  corresponds to the property

$$\forall \mathbf{x}, \mathbf{y} \in D, \quad K(\mathbf{x}, \mathbf{y}) = \langle K(\mathbf{x}, \cdot), K(\cdot, \mathbf{y}) \rangle_{\mathcal{H}_K} \quad (21)$$

for  $\mathcal{H}_K$ .

**Lemma 2:**  $G$  is a closed linear operator from  $\mathcal{H}_C$  onto  $\mathcal{H}_G$ .

*Proof:* Let  $\boldsymbol{\alpha} \in \mathcal{H}_C$  be an arbitrary vector. Then

$$\|G\boldsymbol{\alpha}\|_{\mathcal{H}_G}^2 = \boldsymbol{\alpha}' G M G \boldsymbol{\alpha} = \boldsymbol{\alpha}' G \boldsymbol{\alpha} < \infty \quad (22)$$

is obtained, which implies  $D(G) = \mathcal{H}_C$ , where  $D(G)$  denotes the domain of  $G$ . Let  $\{\boldsymbol{\alpha}_k | k \in \mathbf{N}\}$  be a Cauchy sequence in  $D(G) = \mathcal{H}_C$ . Assume that

$$\lim_{k \rightarrow \infty} \|\boldsymbol{\alpha}_k - \boldsymbol{\alpha}\|_{\mathcal{H}_C}^2 = 0 \quad (23)$$

with  $\boldsymbol{\alpha} \in \mathcal{H}_C$  and

$$\lim_{k \rightarrow \infty} \|G\boldsymbol{\alpha}_k - \boldsymbol{\beta}\|_{\mathcal{H}_G}^2 = 0 \quad (24)$$

<sup>1</sup>Note that it is not required that  $M$  is a 2-inverse [13], which is defined as  $MGM = M$ .

with  $\boldsymbol{\beta} \in \mathcal{H}_G$ . Then

$$\begin{aligned} \|G\boldsymbol{\alpha} - \boldsymbol{\beta}\|_{\mathcal{H}_G}^2 &= \lim_{k \rightarrow \infty} \|G(\boldsymbol{\alpha} - \boldsymbol{\alpha}_k + \boldsymbol{\alpha}_k) - \boldsymbol{\beta}\|_{\mathcal{H}_G}^2 \\ &= \lim_{k \rightarrow \infty} \|G(\boldsymbol{\alpha} - \boldsymbol{\alpha}_k) + (G\boldsymbol{\alpha}_k - \boldsymbol{\beta})\|_{\mathcal{H}_G}^2 \\ &\leq 2 \lim_{k \rightarrow \infty} \|G(\boldsymbol{\alpha} - \boldsymbol{\alpha}_k)\|_{\mathcal{H}_G}^2 \\ &\quad + 2 \lim_{k \rightarrow \infty} \|G\boldsymbol{\alpha}_k - \boldsymbol{\beta}\|_{\mathcal{H}_G}^2 \\ &= 2 \lim_{k \rightarrow \infty} (\boldsymbol{\alpha} - \boldsymbol{\alpha}_k)' G M G (\boldsymbol{\alpha} - \boldsymbol{\alpha}_k) \\ &\quad + 2 \lim_{k \rightarrow \infty} \|G\boldsymbol{\alpha}_k - \boldsymbol{\beta}\|_{\mathcal{H}_G}^2 \\ &= 2 \lim_{k \rightarrow \infty} (\boldsymbol{\alpha} - \boldsymbol{\alpha}_k)' G (\boldsymbol{\alpha} - \boldsymbol{\alpha}_k) \\ &\quad + 2 \lim_{k \rightarrow \infty} \|G\boldsymbol{\alpha}_k - \boldsymbol{\beta}\|_{\mathcal{H}_G}^2 \\ &= 2 \lim_{k \rightarrow \infty} \|\boldsymbol{\alpha} - \boldsymbol{\alpha}_k\|_{\mathcal{H}_C}^2 \\ &\quad + 2 \lim_{k \rightarrow \infty} \|G\boldsymbol{\alpha}_k - \boldsymbol{\beta}\|_{\mathcal{H}_G}^2 = 0 \end{aligned}$$

is obtained, which implies that  $G\boldsymbol{\alpha} = \boldsymbol{\beta}$  and  $G$  is closed. ■

On the basis of the above preliminaries, we have the following theorem.

**Theorem 2:**

$$A \in \mathcal{B}(\mathcal{H}_K, \mathcal{H}_G) \quad (25)$$

where  $\mathcal{B}(\mathcal{H}_K, \mathcal{H}_G)$  denotes the set of bounded linear operators from  $\mathcal{H}_K$  onto  $\mathcal{H}_G$ .

*Proof:* Since  $S$  is a closed linear subspace in  $\mathcal{H}_K$ , any  $f(\cdot) \in \mathcal{H}_K$  is uniquely decomposed as

$$f(\cdot) = f_1(\cdot) + f_2(\cdot), \quad (f_1(\cdot) \in S, \quad f_2(\cdot) \in S^\perp). \quad (26)$$

Since  $f_1(\cdot) \in S$ ,  $f_1(\cdot)$  is represented as

$$f_1(\cdot) = \sum_{i=1}^{\infty} \alpha_i K(\cdot, \mathbf{x}_i) \quad (27)$$

with  $\|f_1(\cdot)\|_{\mathcal{H}_K}^2 = \boldsymbol{\alpha}' G \boldsymbol{\alpha} < \infty$ . Thus

$$\|f(\cdot)\|_{\mathcal{H}_K}^2 = \boldsymbol{\alpha}' G \boldsymbol{\alpha} + C < \infty \quad (28)$$

holds with some non-negative constant  $C$  corresponding to  $\|f_2(\cdot)\|_{\mathcal{H}_K}^2$ . Since  $S^\perp = \mathcal{N}(A)$ ,  $Af_2(\cdot) = 0$  holds. Thus, from Lemma 2

$$\begin{aligned} Af(\cdot) &= Af_1(\cdot) \\ &= \left( \sum_{i=1}^{\infty} [\mathbf{e}_i \otimes K(\cdot, \mathbf{x}_i)] \right) \left( \sum_{j=1}^{\infty} \alpha_j K(\cdot, \mathbf{x}_j) \right) \\ &= \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \mathbf{e}_i K(\mathbf{x}_i, \mathbf{x}_j) \alpha_j = G \boldsymbol{\alpha} \end{aligned}$$

is obtained and well-defined. Therefore, from (20)

$$\begin{aligned} \|Af(\cdot)\|_{\mathcal{H}_G}^2 &= \|Af_1(\cdot)\|_{\mathcal{H}_G}^2 \\ &= \langle Af_1(\cdot), Af_1(\cdot) \rangle_{\mathcal{H}_G} \\ &= \boldsymbol{\alpha}' G M G \boldsymbol{\alpha} = \boldsymbol{\alpha}' G \boldsymbol{\alpha} < \infty \end{aligned}$$

is followed. Thus

$$\begin{aligned} \|A\| &= \sup_{f(\cdot) \in \mathcal{H}_K} \frac{\|Af(\cdot)\|_{\mathcal{H}_G}}{\|f(\cdot)\|_{\mathcal{H}_K}} \\ &= \sup_{\alpha' G \alpha + C < \infty} \left( \frac{\alpha' G \alpha}{\alpha' G \alpha + C} \right)^{\frac{1}{2}} = 1 < \infty \end{aligned} \quad (29)$$

is obtained, which implies  $A \in \mathcal{B}(\mathcal{H}_K, \mathcal{H}_G)$ . ■

According to Theorem 2, it immediately follows that:

$$\begin{aligned} A^* &\in \mathcal{B}(\mathcal{H}_G, \mathcal{H}_K), \\ A^* A &\in \mathcal{B}(\mathcal{H}_K, \mathcal{H}_K). \end{aligned}$$

*Theorem 3:*  $P = A^* A$  is the orthogonal projector onto the closed linear subspace  $S$  in  $\mathcal{H}_K$ .

*Proof:* Let  $f(\mathbf{y}) = \sum_{i=1}^{\infty} \alpha_i K(\mathbf{y}, \mathbf{x}_i)$  be an arbitrary function in  $S$  with respect to  $\mathbf{y}$ , then

$$\begin{aligned} Pf(\mathbf{y}) &= A^* Af(\mathbf{y}) = A^* G \alpha \\ &= \left( \sum_{i=1}^{\infty} K(\mathbf{y}, \mathbf{x}_i) \otimes \mathbf{e}_i \right) G \alpha \\ &= \sum_{i=1}^{\infty} \alpha' G M \mathbf{e}_i K(\mathbf{y}, \mathbf{x}_i) \\ &= \alpha' G M \mathbf{k} \end{aligned}$$

where

$$\mathbf{k} = \sum_{i=1}^{\infty} \mathbf{e}_i K(\mathbf{y}, \mathbf{x}_i) = [K(\mathbf{y}, \mathbf{x}_1), \dots, K(\mathbf{y}, \mathbf{x}_\ell), \dots]' \in \mathbf{R}^\infty.$$

Since  $A \in \mathcal{B}(\mathcal{H}_K, \mathcal{H}_G)$  and  $K(\cdot, \mathbf{y}) \in \mathcal{H}_K$  for any  $\mathbf{y} \in D$

$$\begin{aligned} AK(\cdot, \mathbf{y}) &= \left( \sum_{i=1}^{\infty} [\mathbf{e}_i \otimes K(\cdot, \mathbf{x}_i)] \right) K(\cdot, \mathbf{y}) \\ &= \mathbf{k} \in \mathcal{H}_G \end{aligned}$$

is followed with an arbitrarily fixed  $\mathbf{y} \in D$ . Thus, from (18)

$$\begin{aligned} Pf(\mathbf{y}) &= \alpha' G M \mathbf{k} = \alpha' \mathbf{k} \\ &= \sum_{i=1}^{\infty} \alpha_i K(\mathbf{y}, \mathbf{x}_i) = f(\mathbf{y}) \end{aligned}$$

is obtained for any  $\mathbf{y} \in D$ .

On the other hand, for any  $f(\mathbf{y}) \in S^\perp$

$$Pf(\mathbf{y}) = 0$$

trivially holds for any  $\mathbf{y} \in D$ , since  $S^\perp = \mathcal{N}(A)$ . Thus, it is concluded that  $P = A^* A \in \mathcal{B}(\mathcal{H}_K, \mathcal{H}_K)$  is the orthogonal projector onto the closed linear subspace  $S$  along with  $S^\perp$ . ■

Since  $P = A^* A$  is a bijection from  $S$  onto  $S$ ,  $A$  is an injection at least. Therefore, if  $f_1(\cdot) = \sum_{i=1}^{\infty} \alpha_i K(\cdot, \mathbf{x}_i) \in S$  is arbitrarily fixed and  $Af(\cdot) = G\alpha$  is obtained, then  $f_1(\cdot)$  is the unique element in  $S$  satisfying  $Af(\cdot) = G\alpha$ . Accordingly, any

function  $f(\cdot) \in \mathcal{H}_K$  satisfying  $Af(\cdot) = G\alpha$  can be represented as

$$f(\cdot) = f_1(\cdot) + f_2(\cdot) \quad (30)$$

where  $f_2(\cdot) \in S^\perp$ . Thus, for any  $f(\cdot) \in \mathcal{H}_K$  satisfying  $Af(\cdot) = G\alpha$ , we have

$$\begin{aligned} \|f(\cdot)\|_{\mathcal{H}_K}^2 - \|Af(\cdot)\|_{\mathcal{H}_G}^2 &= \|f(\cdot)\|_{\mathcal{H}_K}^2 - \|G\alpha\|_{\mathcal{H}_G}^2 \\ &= \alpha' G \alpha + C - \alpha' G M G \alpha \\ &= \alpha' G \alpha + C - \alpha' G \alpha \\ &= C \geq 0 \end{aligned}$$

where  $C$  denotes  $\|f_2(\cdot)\|_{\mathcal{H}_K}^2$ ; and  $\|f(\cdot)\|_{\mathcal{H}_K}^2 = \|Af(\cdot)\|_{\mathcal{H}_G}^2$  is achieved when  $f(\cdot) \in S$ , which implies that the norm of  $\mathcal{H}_G$  surely satisfies the property in Theorem 1.

Note that the closed form of  $P$  is written as

$$\begin{aligned} P &= A^* A \\ &= \left( \sum_{j=1}^{\infty} [K(\cdot, \mathbf{x}_j) \otimes \mathbf{e}_j] \right) \left( \sum_{i=1}^{\infty} [\mathbf{e}_i \otimes K(\cdot, \mathbf{x}_i)] \right) \\ &= \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} M_{i,j} [K(\cdot, \mathbf{x}_i) \otimes K(\cdot, \mathbf{x}_j)] \end{aligned} \quad (31)$$

and that of  $Pf(\cdot)$  is written as

$$\begin{aligned} Pf(\cdot) &= \left( \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} M_{i,j} [K(\cdot, \mathbf{x}_i) \otimes K(\cdot, \mathbf{x}_j)] \right) f(\cdot) \\ &= \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} f(\mathbf{x}_j) M_{i,j} K(\cdot, \mathbf{x}_i). \end{aligned} \quad (32)$$

Also note that

$$\|Pf(\cdot)\|_{\mathcal{H}_K}^2 = \mathbf{f}' M \mathbf{f} = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} f(\mathbf{x}_i) M_{i,j} f(\mathbf{x}_j) \quad (33)$$

holds with  $\mathbf{f} = Af(\cdot) (= G\alpha)$ , since

$$\begin{aligned} \|Pf(\cdot)\|_{\mathcal{H}_K}^2 &= \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} \sum_{l=1}^{\infty} f(\mathbf{x}_i) M_{i,j} G_{j,k} M_{k,l} f(\mathbf{x}_l) \\ &= \mathbf{f}' M G M \mathbf{f} = \alpha' G M G M G \alpha = \alpha' G M G \alpha \\ &= \mathbf{f}' M \mathbf{f} \\ &= \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} f(\mathbf{x}_i) M_{i,j} f(\mathbf{x}_j) \end{aligned}$$

is followed.

Since  $P$  is the orthogonal projector onto the closed linear subspace  $S$ ,  $Pf(\cdot)$  gives the optimal approximation of any  $f(\cdot) \in \mathcal{H}_K$  in  $S$ . Thus, the above discussion is an extension of the framework shown in [6]–[8] to infinite sampling points. Fig. 1 illustrates the relationship between operators  $A$ ,  $A^*$ , and  $A^* A$  with RKHS's  $\mathcal{H}_K$  and  $\mathcal{H}_G$ .

In order for (32) to perfectly reconstruct any function  $f(\cdot) \in \mathcal{H}_K$

$$\mathcal{H}_K = S \quad (34)$$

must hold. Thus, what we have to clarify to obtain the sampling theorem in this framework is a necessary and sufficient condition for  $K$  and  $X$  to obtain (34).

#### IV. KERNEL-INDUCED SAMPLING THEOREM

In this section, on the basis of the discussions in the previous section, we give a necessary and sufficient condition for a reproducing kernel and a set of sampling points to perfectly reconstruct any function in the corresponding RKHS. The following theorem is the main result of this paper.

*Theorem 4:*  $\mathcal{H}_K = S$  if and only if

$$K(\mathbf{y}, \mathbf{y}) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} K(\mathbf{y}, \mathbf{x}_j) M_{i,j} K(\mathbf{y}, \mathbf{x}_i) \quad (35)$$

holds for any  $\mathbf{y} \in D$ .

*Proof:* Since  $K(\cdot, \mathbf{y}) \in \mathcal{H}_K$  for any  $\mathbf{y} \in D$ , if  $\mathcal{H}_K = S$  holds

$$K(\cdot, \mathbf{y}) - PK(\cdot, \mathbf{y}) = 0 \quad (36)$$

must hold for any  $\mathbf{y} \in D$  at least.

On the other hand, if we assume that (36) holds, then

$$\begin{aligned} f(\mathbf{y}) &= \langle f(\cdot), K(\cdot, \mathbf{y}) \rangle_{\mathcal{H}_K} \\ &= \langle f(\cdot), PK(\cdot, \mathbf{y}) \rangle_{\mathcal{H}_K} \\ &= \langle f(\cdot), P^* K(\cdot, \mathbf{y}) \rangle_{\mathcal{H}_K} \\ &= \langle Pf(\cdot), K(\cdot, \mathbf{y}) \rangle_{\mathcal{H}_K} \\ &= Pf(\mathbf{y}) \end{aligned}$$

is obtained for any  $f(\cdot) \in \mathcal{H}_K$  and any  $\mathbf{y} \in D$  since  $P$  is an orthogonal projector, which implies  $\mathcal{H}_K = S$ .

It is trivial that (36) is identical to

$$\|K(\cdot, \mathbf{y}) - PK(\cdot, \mathbf{y})\|_{\mathcal{H}_K}^2 = 0 \quad (37)$$

and the Pythagorean theorem and (33) yield

$$\begin{aligned} \|K(\cdot, \mathbf{y}) - PK(\cdot, \mathbf{y})\|_{\mathcal{H}_K}^2 &= \|K(\cdot, \mathbf{y})\|_{\mathcal{H}_K}^2 - \|PK(\cdot, \mathbf{y})\|_{\mathcal{H}_K}^2 \\ &= K(\mathbf{y}, \mathbf{y}) - \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} K(\mathbf{y}, \mathbf{x}_j) M_{i,j} K(\mathbf{y}, \mathbf{x}_i) \\ &= 0 \end{aligned}$$

which concludes the proof.  $\blacksquare$

According to Theorem 4, we can confirm whether a given reproducing kernel and a set of sampling points can perfectly reconstruct all functions in the corresponding RKHS or not by checking (35) for all  $\mathbf{y} \in D$ .

*Some Remarks:* In case of  $|X| < \infty$  (finite sampling points), the above discussions can be applied as it is with  $M = G^+$  (the

Moore–Penrose generalized inverse matrix [14] of  $G$ ), which is reduced to an extension of the framework shown in [6]–[8] for a perfect reconstruction with finite sampling points.

When the set  $\{K(\cdot, \mathbf{x}_i) | i \in \mathbb{N}\}$  is an orthonormal system as in (1), the Gramian matrix  $G$  is reduced to the (infinite dimensional) identity matrix  $I$ . In this case, the matrix  $M$  that specifies the metric in  $\mathcal{H}_G$  is also reduced to  $I$ , which implies that  $\mathcal{H}_G$  is identical to  $\ell^2$ , the Hilbert space of square summable vectors.

Note that when  $G$ , with infinite sampling points, is not the identity matrix,  $\mathcal{H}_G$  is not always a subset of  $\ell^2$ . In fact, an extreme example  $X = \{\mathbf{x}_i = \mathbf{c} | i \in \{1, \dots, n\}, \mathbf{c} \in D\}$  yields

$$\mathbf{f}_n = Af(\cdot) = [f(\mathbf{c}), \dots, f(\mathbf{c})]' \in \mathbf{R}^n$$

and

$$\|\lim_{n \rightarrow \infty} \mathbf{f}_n\|_{\ell^2}^2 = \lim_{n \rightarrow \infty} \mathbf{f}_n' \mathbf{f}_n = \infty \quad (38)$$

when  $f(\mathbf{c}) \neq 0$ , which implies  $\lim_{n \rightarrow \infty} \mathbf{f}_n \notin \ell^2$ . On the other hand, the Gramian matrix of  $K$  corresponding to  $X$  is reduced to

$$G_n = \begin{bmatrix} K(\mathbf{c}, \mathbf{c}) & \cdots & K(\mathbf{c}, \mathbf{c}) \\ \vdots & \ddots & \vdots \\ K(\mathbf{c}, \mathbf{c}) & \cdots & K(\mathbf{c}, \mathbf{c}) \end{bmatrix} \in \mathbf{R}^{n \times n} \quad (39)$$

and its Moore–Penrose generalized inverse matrix is given as

$$M_n = \frac{1}{n^2 K(\mathbf{c}, \mathbf{c})} \begin{bmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{bmatrix} \in \mathbf{R}^{n \times n} \quad (40)$$

when  $K(\mathbf{c}, \mathbf{c}) \neq 0$ , then

$$\|\lim_{n \rightarrow \infty} \mathbf{f}_n\|_{\mathcal{H}_G}^2 = \lim_{n \rightarrow \infty} \mathbf{f}_n' M_n \mathbf{f}_n = \frac{f(\mathbf{c})^2}{K(\mathbf{c}, \mathbf{c})} < \infty \quad (41)$$

holds, which implies  $\lim_{n \rightarrow \infty} \mathbf{f}_n \in \mathcal{H}_G$ . When  $K(\mathbf{c}, \mathbf{c}) = 0$ ,  $M_n \in \mathbf{R}^{n \times n}$  is reduced to the zero matrix, which trivially implies that  $\lim_{n \rightarrow \infty} \|\mathbf{f}_n\|_{\mathcal{H}_G}^2 = 0 < \infty$  and  $\lim_{n \rightarrow \infty} \mathbf{f}_n \in \mathcal{H}_G$ . The key idea that makes our framework consistent for all cases is adopting  $\mathcal{H}_G$  as the range space of the sampling operator  $A$ .

Since  $K(\cdot, \mathbf{x}) \in \mathcal{H}_K$  for any  $\mathbf{x} \in D$ , it is obvious that (36) holds when  $\mathcal{H}_K = S$ . In fact, formulas similar to (36) is given in [4] for some cases. However, the sufficiency of (36) [or (35)] for perfect reconstruction with infinite sampling points is never mentioned before.<sup>2</sup> Thus, the statement for its sufficiency in Theorem 4, proved by incorporating the range space  $\mathcal{H}_G$  of the sampling operator, is our main contribution.

#### V. ERROR ANALYSIS FOR INCOMPLETE SAMPLING POINTS

In this section, we give an error analysis for a reconstructed function by the orthogonal projector  $P$  with incomplete sampling points, which is the case of  $S \neq \mathcal{H}_K$ .

When  $X$  is incomplete for perfect reconstruction of a function in  $\mathcal{H}_K$ , (35) and (36) do not hold. Thus, at an arbitrarily fixed

<sup>2</sup>Our main result given as Theorem 4 can be regarded as an extension of [4, Prop. 4.2] with infinite sampling points. In fact, (35) is quite similar to the norm-representation of [4, Eq. (4.6)] with infinite sampling points.

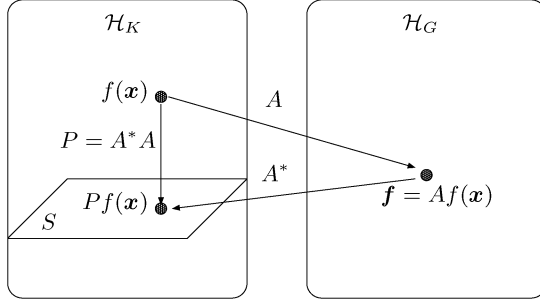


Fig. 1. Relationship between  $A$ ,  $A^*$ , and  $A^*A$  with the RKHS's  $\mathcal{H}_K$  and  $\mathcal{H}_G$ .

$\mathbf{x} \in D$ , the absolute difference between  $f(\mathbf{x})$  and  $Pf(\mathbf{x})$  is reduced to

$$\begin{aligned} |f(\mathbf{x}) - Pf(\mathbf{x})| &= |\langle f(\cdot), K(\cdot, \mathbf{x}) \rangle_{\mathcal{H}_K} - \langle Pf(\cdot), K(\cdot, \mathbf{x}) \rangle_{\mathcal{H}_K}| \\ &= |\langle f(\cdot), K(\cdot, \mathbf{x}) \rangle_{\mathcal{H}_K} - \langle f(\cdot), PK(\cdot, \mathbf{x}) \rangle_{\mathcal{H}_K}| \\ &= |\langle f(\cdot), K(\cdot, \mathbf{x}) - PK(\cdot, \mathbf{x}) \rangle_{\mathcal{H}_K}| \\ &\leq \|f(\cdot)\|_{\mathcal{H}_K} \|K(\cdot, \mathbf{x}) - PK(\cdot, \mathbf{x})\|_{\mathcal{H}_K} \\ &= \|f(\cdot)\|_{\mathcal{H}_K} E_{K,X}(\mathbf{x}) \end{aligned}$$

by applying the Schwarz's inequality, where

$$E_{K,X}(\mathbf{x}) = \left( K(\mathbf{x}, \mathbf{x}) - \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} K(\mathbf{x}, \mathbf{x}_j) M_{i,j} K(\mathbf{x}, \mathbf{x}_i) \right)^{1/2}. \quad (42)$$

Thus, it is concluded that the absolute reconstruction error at the point  $\mathbf{x}$  is bounded by the value proportional to  $\|f(\cdot)\|_{\mathcal{H}_K}$  and  $E_{K,X}(\mathbf{x})$ . Accordingly, when the target function  $f(\cdot)$  is normalized by its norm, the absolute reconstruction error is bounded by  $E_{K,X}(\mathbf{x})$  that only depends on the kernel and the set of sampling points. This fact implies that we can identify the point  $\mathbf{x} \in D$ , where the absolute reconstruction error tends to be large, without information of the target function  $f(\cdot)$ , which may be useful not only for applications in signal processing but also for a model selection in the kernel-based learning theory (see [15] for instance).

## VI. EXAMPLES

In this section, we show four examples for our results. Examples VI-A and VI-B give another proof of Shannon's sampling theorem and a sampling theorem for the RKHS corresponding to a polynomial kernel, respectively. Example VI-C reveals that the Sobolev space does not have a sampling theorem with equally spaced sampling points whose interval is larger than 0. Example VI-D gives an error analysis with an incomplete set of sampling points.

### A. Shannon's Sampling Theorem

It is mentioned in [4] that sinc function written as

$$K(x, y) = \frac{\sin \pi(x - y)}{\pi(x - y)}, \quad x, y \in \mathbf{R} \quad (43)$$

is the reproducing kernel of the class of  $\pi$ -bandlimited functions with finite energy. Let  $X = \{\dots, -1, 0, 1, \dots\}$  be the set

of sampling points with the Nyquist interval for  $\pi$ -bandlimited functions. Note that the Gramian matrix for these  $K$  and  $X$  is reduced to the identity operator and the matrix  $M$  is also reduced to the identity operator. It is trivial that  $K(y, y) = 1$  for any  $y \in \mathbf{R}$  in the left-hand side of (35). The right-hand side of (35) is reduced to

$$\sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} K(y, i) M_{i,j} K(y, j) = \sum_{i=-\infty}^{\infty} \left( \frac{\sin \pi(y - i)}{\pi(y - i)} \right)^2. \quad (44)$$

When  $y \in X$ , it is easy to show that (44) is equal to 1. On the other hand, when  $y \notin X$

$$\begin{aligned} &\sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} K(y, i) M_{i,j} K(y, j) \\ &= \sum_{i=-\infty}^{\infty} \left( \frac{\sin \pi(y - i)}{\pi(y - i)} \right)^2 \\ &= \sum_{i=-\infty}^{\infty} \left( \frac{\sin \pi y \cos \pi i - \cos \pi y \sin \pi i}{\pi(y - i)} \right)^2 \\ &= \sum_{i=-\infty}^{\infty} \left( \frac{(-1)^i \sin \pi y}{\pi(y - i)} \right)^2 \\ &= \sin^2 \pi y \sum_{i=-\infty}^{\infty} \left( \frac{1}{\pi(y - i)} \right)^2 \\ &= \sin^2 \pi y \operatorname{cosec}^2 \pi y = 1 \end{aligned}$$

holds for any  $y \notin X$ . Thus, it is concluded that (35) holds for any  $y \in \mathbf{R}$ , which gives another proof of Shannon's sampling theorem.

### B. Sampling Theorem for Polynomial Kernels

Let

$$K_P^{(2)}(x, y) = (1 + xy)^2, \quad x, y \in [-1, 1] \subset \mathbf{R} \quad (45)$$

be a polynomial kernel of degree 2 defined on  $[-1, 1]$ . The Mercer expansion [10], [11] of  $K_P^{(2)}(x, y)$  is given as

$$K_P^{(2)}(x, y) = \sum_{i=1}^3 \lambda_i \phi_i(x) \phi_i(y) \quad (46)$$

where

$$\lambda_1 = \frac{18 + 2\sqrt{61}}{15}, \quad \lambda_2 = \frac{4}{3}, \quad \lambda_3 = \frac{18 - 2\sqrt{61}}{15}$$

and

$$\phi_1(x) = \frac{\sqrt{15}}{4\sqrt{183 - 23\sqrt{61}}} \left( (3\sqrt{61} - 23)x^2 + (9 - \sqrt{61}) \right)$$

$$\phi_2(x) = \frac{\sqrt{3}}{\sqrt{2}} x$$

$$\phi_3(x) = \frac{\sqrt{3}}{2\sqrt{183 - 23\sqrt{61}}} \left( 5x^2 + (6 - \sqrt{61}) \right)$$

are orthonormal system in  $L^2([-1, 1])$ . Thus, it is suggested that any function in the corresponding RKHS, denoted by  $\mathcal{H}_{K_P^{(2)}}$ , can be perfectly reconstructed by three sampling points.

Let  $X = \{x_1, x_2, x_3 | x_i \in [-1, 1]\}$  be an arbitrary set of sampling points with  $x_i \neq x_j$  for  $i \neq j$ . It is trivial that  $K(y, y) = (1 + y^2)^2$  for any  $y \in [-1, 1]$  in the left-hand side of (35). The Gramian matrix  $G$  is given as

$$G = \begin{bmatrix} (1 + x_1^2)^2 & (1 + x_1x_2)^2 & (1 + x_1x_3)^2 \\ (1 + x_2x_1)^2 & (1 + x_2^2)^2 & (1 + x_2x_3)^2 \\ (1 + x_3x_1)^2 & (1 + x_3x_2)^2 & (1 + x_3^2)^2 \end{bmatrix} \quad (47)$$

and its inverse matrix  $M = G^{-1}$  is reduced to

$$M = \begin{bmatrix} m_1 & m_2 & m_3 \\ m_2 & m_4 & m_5 \\ m_3 & m_5 & m_6 \end{bmatrix} \quad (48)$$

where

$$\begin{aligned} m_1 &= \frac{2x_2^2x_3^2 + x_3^2 + 2x_2x_3 + x_2^2 + 2}{2(x_2 - x_1)^2(x_3 - x_1)^2} \\ m_2 &= -\frac{2x_1x_2x_3^2 + x_3^2 + x_2x_3 + x_1x_3 + x_1x_2 + 2}{2(x_2 - x_1)^2(x_3 - x_1)(x_3 - x_2)} \\ m_3 &= \frac{2x_1x_2^2x_3 + x_2x_3 + x_1x_3 + x_2^2 + x_1x_2 + 2}{2(x_2 - x_1)(x_3 - x_1)^2(x_3 - x_2)} \\ m_4 &= \frac{2x_1^2x_3^2 + x_3^2 + 2x_1x_3 + x_1^2 + 2}{2(x_2 - x_1)^2(x_3 - x_2)^2} \\ m_5 &= -\frac{2x_1^2x_2x_3 + x_2x_3 + x_1x_3 + x_1x_2 + x_1^2 + 2}{2(x_2 - x_1)(x_3 - x_1)(x_3 - x_2)^2} \\ m_6 &= \frac{2x_1^2x_2^2 + x_2^2 + 2x_1x_2 + x_1^2 + 2}{2(x_3 - x_1)^2(x_3 - x_2)^2}. \end{aligned}$$

Thus, the right-hand side of (35) is reduced to

$$\begin{aligned} &\sum_{i=1}^3 \sum_{j=1}^3 K_P^{(2)}(y, x_i) M_{i,j} K_P^{(2)}(y, x_j) \\ &= \begin{bmatrix} (1 + x_1y)^2 & (1 + x_2y)^2 & (1 + x_3y)^2 \end{bmatrix} M \\ &\quad \times \begin{bmatrix} (1 + x_1y)^2 & (1 + x_2y)^2 & (1 + x_3y)^2 \end{bmatrix}' \\ &= (1 + y^2)^2 \end{aligned}$$

for any  $y \in [-1, 1]$ . Therefore, it is concluded that any function in  $\mathcal{H}_{K_P^{(2)}}$  can be perfectly reconstructed with arbitrary (but different) three sampling points. This fact is consistent with the suggestion by the Mercer expansion given above.

Note that these results for  $K_P^{(2)}$  can be easily extended to a polynomial kernel with higher degree.

### C. Sampling Theorem for Sobolev Spaces

In [4], sampling theorem for subspaces of the Sobolev space  $\mathcal{H}^1$  is obtained. Here, we discuss the sampling theorem for  $\mathcal{H}^1$  itself based on Theorem 4. The reproducing kernel of  $\mathcal{H}^1$  is given as

$$K_s^{(\alpha)}(x, y) = \frac{1}{2} \exp(-\alpha|x - y|), \quad x, y \in \mathbf{R} \quad (49)$$

with  $\alpha > 0$ . Let  $X = \{\dots, -1, 0, 1, \dots\}$  be the set of sampling points and let  $a = \exp(-\alpha)$ , then the Gramian matrix of  $K_s^{(\alpha)}$  and  $X$  is written as

$$G = \frac{1}{2} \begin{bmatrix} \ddots & \ddots & \ddots & \ddots & \ddots \\ \ddots & 1 & a & a^2 & \ddots \\ \ddots & a & 1 & a & \ddots \\ \ddots & a^2 & a & 1 & \ddots \\ \ddots & \ddots & \ddots & \ddots & \ddots \end{bmatrix}. \quad (50)$$

As shown in [4],  $G$  is invertible and its inverse is given as

$$M = \frac{2}{1 - a^2} \begin{bmatrix} \ddots & \ddots & \ddots & \ddots & \ddots \\ \ddots & 1 + a^2 & -a & 0 & \ddots \\ \ddots & -a & 1 + a^2 & -a & \ddots \\ \ddots & 0 & -a & 1 + a^2 & \ddots \\ \ddots & \ddots & \ddots & \ddots & \ddots \end{bmatrix}. \quad (51)$$

Thus, the right-hand side of (35) can be written as

$$\begin{aligned} &\sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} K_s^{(\alpha)}(y, x_j) M_{i,j} K_s^{(\alpha)}(y, x_i) \\ &= \frac{1}{4} \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} a^{|y-i|} M_{i,j} a^{|y-j|} \\ &= \frac{1}{2(1 - a^2)} \sum_{i=-\infty}^{\infty} a^{|y-i|} \\ &\quad \times \left( (1 + a^2) a^{|y-i|} - a^{|y-(i+1)|+1} - a^{|y-(i-1)|+1} \right) \\ &= \frac{1}{2(1 - a^2)} \sum_{i=-\infty}^{\infty} \left( (1 + a^2) a^{2|y-i|} - a^{|y-i|+|y-(i+1)|+1} \right. \\ &\quad \left. - a^{|y-i|+|y-(i-1)|+1} \right). \quad (52) \end{aligned}$$

Without loss of generality, we can assume  $0 \leq y \leq 1$ . Thus, (52) is reduced to

$$\begin{aligned} &\frac{1}{2(1 - a^2)} \left( (1 + a^2) \frac{a^{2y} + a^{2-2y}}{1 - a^2} - \right. \\ &\quad \left. 2a^2 \left( 1 + \frac{1}{1 - a^2} (a^{2y} + a^{2-2y}) \right) \right) = \frac{a^{2y} + a^{2-2y} - 2a^2}{2(1 - a^2)}. \quad (53) \end{aligned}$$

On the other hand, it is trivial that the left-hand side of (35) is equal to  $1/2$ . Thus, in order to obtain the sampling theorem for  $\mathcal{H}^1$  with  $X$

$$\frac{a^{2y} + a^{2-2y} - 2a^2}{2(1 - a^2)} = \frac{1}{2} \quad (54)$$

must hold for any  $y \in [0, 1]$ . Eq. (54) is identical to

$$a^y + a^{1-y} = 1 + a. \quad (55)$$

It is obvious that when  $y \in \{0, 1\}$ , (55) holds, which means that function values at the sampling points can be perfectly reconstructed. However, when  $y \notin \{0, 1\}$ , (55) never holds with  $\alpha > 0$  ( $a < 1$ ). Note that reconstruction of functions



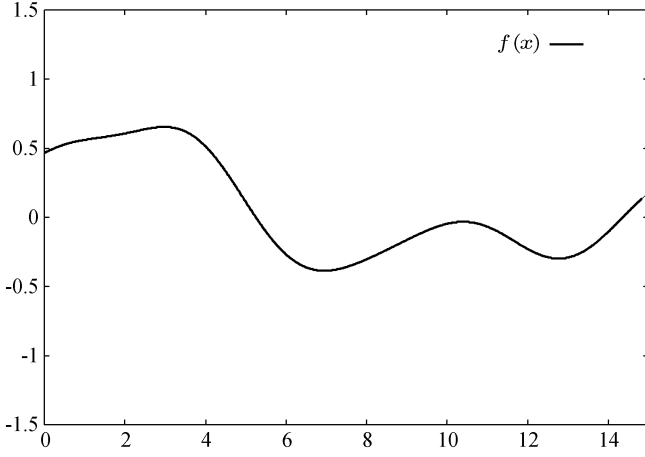


Fig. 2. An example of a target function.

in the RKHS corresponding to  $K_s^{(\alpha)}$  by  $X$  is identical to reconstruction of functions in the RKHS corresponding to  $K_s^{(1)}$  by  $X^{(\alpha)} = \{\dots, -\alpha, 0, \alpha, \dots\}$ . Accordingly, it is concluded that the RKHS corresponding to  $K_s^{(1)}$  does not have sampling theorem by equally spaced sampling points whose interval is larger than 0.

#### D. Incomplete Sampling Points for Gaussian Kernel

Consider the Gaussian kernel, written as

$$K_g(x, y) = \exp\left(-\frac{(x-y)^2}{2\sigma^2}\right), \quad x, y \in \mathbf{R} \quad (56)$$

with the kernel parameter  $\sigma^2 = 2$ , and the corresponding RKHS  $\mathcal{H}_g$ . Let

$$f(\cdot) = \sum_{i=1}^{300} c_i K_g(\cdot, p_i) \quad (57)$$

be the target function in  $\mathcal{H}_g$ , where  $p_i \in [0, 15]$  and  $c_i \in \mathbf{R}$  are randomly generated with the constraint  $\|f(\cdot)\|_{\mathcal{H}_g}^2 = 1$ . Fig. 2 shows the instance of (57) used in the following contents. Let

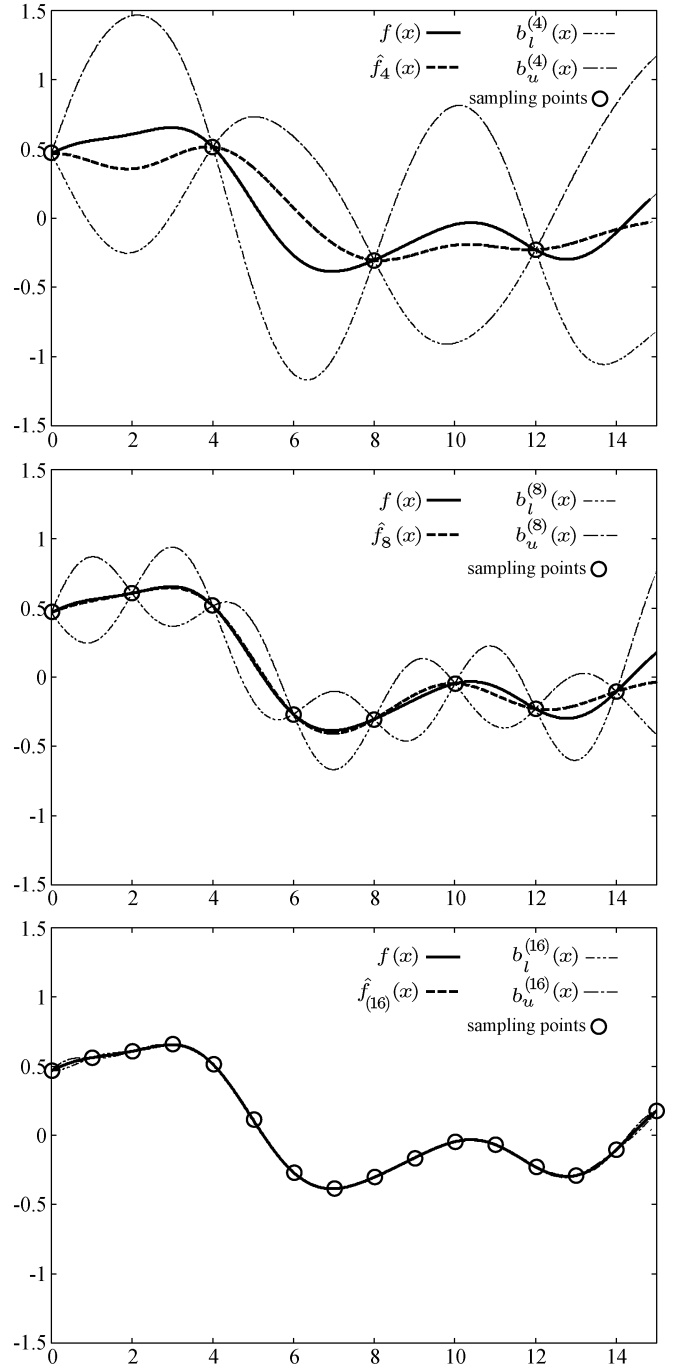
$$\begin{aligned} X_4 &= \{0, 4, 8, 12\}, & |X_4| &= 4 \\ X_8 &= \{0, 2, 4, \dots, 14\}, & |X_8| &= 8 \\ X_{16} &= \{0, 1, 2, \dots, 15\}, & |X_{16}| &= 16 \end{aligned}$$

be the examples of incomplete sampling points for perfect reconstruction of functions in  $\mathcal{H}_g$  and let  $\hat{f}_n(\cdot)$  be the optimal reconstructed function with sampling points  $X_n$  ( $n = 4, 8, 16$ ) by the orthogonal projection (32). According to the analysis in Section V

$$b_l^{(n)}(x) \leq \hat{f}_n(x) \leq b_u^{(n)}(x) \quad (58)$$

must hold with  $n = 4, 8, 16$ , where

$$\begin{aligned} b_u^{(n)}(x) &= f(x) + E_{K_g, X_n}(x) \\ b_l^{(n)}(x) &= f(x) - E_{K_g, X_n}(x) \end{aligned}$$

Fig. 3. The target function, the reconstructed function, and the lower and the upper bounds with  $X_4$  (upper),  $X_8$  (middle), and  $X_{16}$  (lower), respectively.

are the upper and the lower bound functions obtained in Section V. Fig. 3 shows the graphs of  $f(x)$ ,  $\hat{f}_n(x)$ ,  $b_u^{(n)}(x)$ , and  $b_l^{(n)}(x)$  with  $n = 4, 8, 16$ , respectively, which supports the validity of the contents in Section V since it is confirmed numerically that (58) is satisfied in all cases. According to Fig. 3,  $X_{16}$  seems near complete for functions in  $\mathcal{H}_g$  with  $[0, 15]$ , while  $X_4$  seems quite insufficient for function reconstruction in this case. These results also reveal the overestimation of the error bound due to the approximation by Schwarz's

inequality. Again note that  $E_{K,X}(x)$  does not depend on the target function.

## VII. CONCLUSION

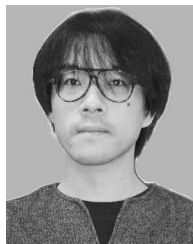
In this paper, we gave a necessary and sufficient condition for the pair of a reproducing kernel and a set of sampling points to perfectly reconstruct any function in the reproducing kernel Hilbert space corresponding to the adopted kernel. We also gave an error analysis of the optimal approximation given by incomplete sampling points. On the basis of our results, we showed another proof of Shannon's sampling theorem and introduced a sampling theorem for the reproducing kernel Hilbert spaces corresponding to a polynomial kernel; and also showed that Sobolev spaces do not have a sampling theorem by equally spaced sampling points. An error analysis with incomplete sampling points for the reproducing kernel Hilbert space corresponding to the Gaussian kernel was also given.

## ACKNOWLEDGMENT

The authors are most grateful to anonymous reviewers for their fruitful comments that improve the quality of this paper.

## REFERENCES

- [1] C. E. Shannon, "A mathematical theory of communication," *Mobile Comput. Commun. Rev.*, vol. 5, no. 1, pp. 3–55, 2001.
- [2] M. Unser, "Sampling—50 years after Shannon," *Proc. IEEE*, vol. 88, no. 4, pp. 569–587, 2000.
- [3] A. I. Zayed, *Advances in Shannon's Sampling Theory*. Boca Raton, FL: CRC Press, 1993.
- [4] M. Z. Nashed and G. G. Walter, "General sampling theorem for functions in reproducing kernel Hilbert space," *Math. Control, Signals, and Syst.*, vol. 4, no. 4, pp. 363–390, 1991.
- [5] G. G. Walter, "A sampling theorem for wavelet subspaces," *IEEE Trans. Inf. Theory*, vol. 38, pp. 881–884, Mar. 1992.
- [6] H. Ogawa, "What can we see behind sampling theorems?," *IEICE Trans. Fundamentals*, vol. E92-A, no. 3, pp. 688–707, 2009.
- [7] H. Ogawa, "Neural networks and generalization ability," (in Japanese) *IEICE Tech. Rep.*, vol. NC95-8, pp. 57–64, 1995.
- [8] A. Hirabayashi, H. Ogawa, and Y. Yamashita, "Admissibility of memorization learning with respect to projection learning in the presence of noise," *IEICE Trans. Inf. Syst.*, vol. E82-D, no. 2, pp. 488–496, 1999.
- [9] N. Aronszajn, "Theory of reproducing kernels," *Trans. Amer. Math. Soc.*, vol. 68, no. 3, pp. 337–404, 1950.
- [10] J. Mercer, "Functions of positive and negative type and their connection with the theory of integral equations," *Trans. London Philosophical Soc.*, vol. A, no. 209, pp. 415–446, 1909.
- [11] S. Saitoh, *Integral Transforms, Reproducing Kernels and Their Applications*. London, U.K.: Addison Wesley Longman Ltd., 1997.
- [12] R. Schatten, *Norm Ideals of Completely Continuous Operators*. Berlin, Germany: Springer-Verlag, 1960.
- [13] A. Ben-Israel and T. N. E. Greville, *Generalized Inverses: Theory and Applications (Second Edition)*. Berlin, Germany: Springer-Verlag, 2003.
- [14] C. R. Rao and S. K. Mitra, *Generalized Inverse of Matrices and Its Applications*. New York: Wiley, 1971.
- [15] K. Müller, S. Mika, G. Ratsch, K. Tsuda, and B. Scholkopf, "An introduction to kernel-based learning algorithms," *IEEE Trans. Neural Networks*, vol. 12, pp. 181–201, Mar. 2001.



**Akira Tanaka** received the D.E. degree from Hokkaido University, Sapporo, Japan, in 2000.

He is with the Graduate School of Information Science and Technology, Hokkaido University. His research interests include image processing, acoustic signal processing, and machine learning.



**Hideyuki Imai** received the D.E. degree from Hokkaido University, Sapporo, Japan, in 1999.

He is with the Graduate School of Information Science and Technology, Hokkaido University. His research interests include statistical inference.



**Masaaki Miyakoshi** received the D.E. degree from Hokkaido University, Sapporo, Japan, in 1985.

He is with the Graduate School of Information Science and Technology, Hokkaido University. His research interests include fuzzy theory.