Sampling Signals from a Union of Subspaces

Yue M. Lu and Minh N. Do

I. INTRODUCTION

Our entire digital revolution depends on the sampling process, which converts continuous-domain real-life signals to sequences of discrete numbers that can be processed and stored by digital devices. Figure 1 shows the basic setup of the sampling problem. The underlying signal x is measured through an acquisition or sampling process A, which produces a discrete sequence of measurements $b = Ax$. A classical setting is the sampling of bandlimited functions [1], where the acquisition step A is often implemented as lowpass anti-aliasing filtering followed by uniform pointwise sampling. The last step of the sampling problem is the reconstruction process $\hat{x} = R(b)$, which aims to re-create a continuous-domain signal $\hat{x}$ from b.

Converting continuous-domain signals into discrete sequences of numbers, the sampling process involves a dramatic reduction of information. Therefore, only when we have some additional *a priori* information about the original signal x, can we possibly hope that the reconstruction $\hat{x}$ from the samples be a faithful or even perfect copy of x.

One of the fundamental assumptions in several classical sampling theorems [1], [2] is that the class $\mathcal{X}$ of signals to be sampled is a single linear vector space. For example, the celebrated Kotelnikov-Shannon-Whittaker sampling theorem deals with bandlimited signals; while for a given spectral support $\mathcal{B}$, it is easy to verify that all signals bandlimited to $\mathcal{B}$ form a single vector space. Similarly, in many generalizations of the sampling theorem (see, for example, [2], [3] and the references therein), the signal class $\mathcal{X}$ is often considered to be a shift-invariant space, e.g., the space of uniform splines.

The single linear vector space assumption is widely used in modeling the signal classes, mainly due to its simplicity and mathematical tractability. However, as we will illustrate with the following examples, in many situations, a *union of subspaces* can be a more appropriate model for the underlying signals of interest.
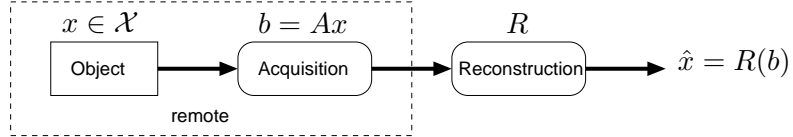
Fig. 1. The general setup for the sampling (or sensing) problem.

Example 1 (Overlapping echoes): In many applications such as geophysics, radar, and wireless communications, the observed signals are often in the form [4]

$$x(t) = \sum_{k=1}^{K} c_k \, \phi(t - t_k), \tag{1}$$

where the pulse shape $\phi(t)$ is assumed to be known; while the delays $\{t_k\}_{k=1}^{K}$ and amplitudes $\{c_k\}_{k=1}^{K}$ are unknown. Note that the signals in (1) are only specified by $2K$ parameters (i.e. K delays and K amplitude). So, intuitively, we should be able to capture the signals by using only $2K$ discrete measurements, which can be much more efficient than sampling at the Nyquist rate.

In fact, this has been a major motivation behind the recent work on sampling signals with *finite rate of innovation* (FRI) [5], [6], which studies several specific classes of signals in the form of (1), including streams of Diracs and piecewise polynomials. One of the main results of the FRI sampling work is to demonstrate that these classes of non-bandlimited signals can be uniformly sampled at a low given rate (i.e. the "rate of innovation"), and then perfectly reconstructed by using efficient algorithms based on annihilating filter techniques.

Observe that, once we fix the K delay values $\{t_k\}$ in (1) but let the amplitude $\{c_k\}$ change freely, then the resulting signals live in a linear subspace of dimension up to K, spanned by $\{\phi(t - t_k)\}_{k=1}^{K}$. Therefore, the entire signal class $\mathcal{X}$ can be modeled as a union of subspaces, each of which corresponds to a set of possible delays $\{t_k\}_{k=1}^{K}$.

Example 2 (Signals with unknown spectral support): Consider the class of continuous multi-band signals $\mathcal{X}^{\mathrm{MB}}$, whose Fourier transforms only occupy a known fraction – but at *unknown* locations – on a spectral band $[f_{\min}, f_{\max}]$. To simplify the problem, we can divide the spectral span $\mathcal{F} = [f_{\min}, f_{\max}]$ into L equally spaced spectral cells $\{\mathcal{C}_i\}_{i=0}^{L-1}$, and assume that the signals to be sampled only have nonzero frequency values in at most K spectral cells (with $K \ll L$), though we do not know the exact locations of them.

If we know exactly which K cells contain nonzero frequency values, then we have a classical

sampling problem, for which the minimum sampling rate was given by Landau [7] and is equal to $(f_{\max} - f_{\min})\frac{K}{L}$. However, the difficulty here is that we do not have *a priori* information on the spectral support, except for its occupancy ratio. We can still sample the signals at the Nyquist sampling rate, which equals $f_{\max} - f_{\min}$; but this rate can be much higher than the Landau lower bound, especially when the signal spectrum is sparse. In a series of pioneering papers [8], [9], Bresler *et al.* showed that the class of signals $\mathcal{X}^{\mathrm{MB}}$ can be sampled and reconstructed by a universal non-adaptive sampling scheme, whose rate can be much lower than the Nyquist rate.

We see that the signal class $\mathcal{X}^{\mathrm{MB}}$ is not a single linear subspace, because the summation of two multiband signals with different spectral supports can occupy $2K$ spectral cells and thus does not belong to $\mathcal{X}^{\mathrm{MB}}$. However, for a fixed set of spectral cells $\{\mathcal{C}_{i_1}, \ldots, \mathcal{C}_{i_K}\}$, the signals occupying these spectral cells live in the subspace of functions bandlimited to $\bigcup_{k=1}^{K} \mathcal{C}_{i_k}$. Therefore, with all possible configurations of the spectral supports, the signal class can be characterized by a union of subspaces, each specified by a particular selection of K cells.

Example 3 (Sparse representation): Many practical signals (e.g. audio and images) can be well-approximated by only a small number of basis elements from a fixed dictionary $\{\phi_k\}_{k=1}^{\infty}$ (e.g. a wavelets basis). In these cases, the output signal is a K-term approximation, written as

$$\hat{f}_K = \sum_{k \in I} c_k \phi_k, \tag{2}$$

where I is an index set of K selected basis functions or atoms. Such signals are studied in a recent breakthrough in mathematics under the name *compressed sensing* or *compressive sampling* [10]–[12], for the case when the underlying dictionary $\{\phi_k\}$ is a *finite* set. The pioneering contribution of the compressed sensing/compresive sampling work shows that finite-length sparse signals as in (2) can be fully recovered from a small number of linear, non-adaptive, and random measurements.

We observe once again that the set of signals that can be represented by K terms from a given basis or dictionary also constitutes a union of subspaces, with each subspace indexed by a set I.

Through the three examples given above, we have shown that the traditional single-space (e.g. bandlimited) sampling theorems can be inefficient for several important classes of signals. This limitation has been previously recognized and addressed in various new sampling schemes, including the spectrum-blind sampling [8], finite-rate of innovation sampling [5], and compressed sensing/compressive sampling [10]–[12]. The focus of this article is to provide a new perspective

for the extension of the sampling theory, by considering signals from a *union of subspaces* instead of a *single space*. This quite general model not only covers all the signals discussed in the above examples, but also incorporates, as we will show in later discussions, new classes of signals that have not been previously considered in related work.

The presentation of this article goes as follows. We first describe a linear sampling framework for signals from a union of subspaces and provide a corresponding geometric interpretation. We then present some key results on conditions of the invertible and stable sampling operators, with applications in determining the minimum sampling requirement. These conditions offer fundamental bounds on sampling signals from a union of subspaces. We finish our discussions by describing several applications of the theoretical results obtained under the proposed sampling framework.

II. A Linear Sampling Framework for Signals from a Union of Subspaces

A. Sampling Framework and Problem Formulation

The examples given in the previous section motivate us to consider the following abstract definition for many classes of signals of interest.

First, let $\mathcal{H}$ be an ambient Hilbert space that our signals live in. We assume the signals of interest span a fixed *union of subspaces* that is defined as

$$\mathcal{X} = \bigcup_{\gamma \in \Gamma} \mathcal{S}_\gamma, \quad \text{where } \mathcal{S}_\gamma \text{ are subspaces of } \mathcal{H} \text{ and } \Gamma \text{ is an index set.} \tag{3}$$

In other words, a signal $x \in \mathcal{X}$ if and only if there is some $\gamma_0 \in \Gamma$ such that $x \in S_{\gamma_0}$.

We consider a general sampling framework in which the input signal $x \in \mathcal{X}$ is sampled via a *linear mapping* A into a sequence of numbers $\{(Ax)_n\}_{n \in \Lambda}$, where the nth sample value $(Ax)_n$ is the inner product of x with a sampling vector $\psi_n \in \mathcal{H}$, i.e.,

$$(Ax)_n = \langle x, \psi_n \rangle_{\mathcal{H}}. \tag{4}$$

Note that in (4), ψ_n resembles the point spreading function of the nth measurement device.

Given a class of signals defined as a union of subspaces in (3), it is natural to pursue the following standard questions related to sampling and reconstruction.

1) When is each object $x \in \mathcal{X}$ *uniquely represented by its sampling data* $\{\langle x, \psi_n \rangle\}_{n \in \Lambda}$?

2) What is the *minimum sampling requirement* for a signal class $\mathcal{X}$?

3) What are the *optimal sampling functions* $\{\psi_n\}_{n \in \Lambda}$?

4) What are *efficient algorithms to reconstruct a signal* $x \in \mathcal{X}$ from its sampling data $\{\langle x, \psi_n \rangle\}_{n \in \Lambda}$?

5) How *stable is the reconstruction* in the presence of noise and model mismatch (e.g. the signals may only approximately live in a union of subspaces)?

As discussed in Section I, for some particular subclasses of the signals that can be modeled as coming from a union of subspaces, the above questions have already been studied and answered (to various degrees) in related work such as the spectrum-blind sampling [8] (for multiband signals), finite rate of innovation sampling [5] (for streams of Diracs[1], piecewise polynomials, etc.), and compressed sensing/compressive sampling [10]–[12] (for finite-length sparse or compressible discrete signals).

The main contribution of this article is to address the first two questions listed above in their most general setting, with the only assumption that the signals come from some known union of subspaces. In particular, instead of focusing on finite-lengh discrete cases, our discussions allow the signals to be much more general, for example, as continuous-domain functions coming from infinite-dimensional shift-invariant subspaces.

B. A Geometrical Viewpoint

There is a nice geometric interpretation for the questions posed above. First we recall that, in the Hilbert space $\mathcal{H}$, knowing $\{\langle x, \psi_n \rangle\}_{n \in \Lambda}$ is equivalent to knowing the projection $\mathbf{P}_{\mathcal{S}} x$ of x onto the subspace $\mathcal{S} = \overline{\mathrm{span}} \{\psi_n\}_{n \in \Lambda}$. We call $\mathcal{S}$ a *representation subspace*. Geometrically, we can think of the linear sampling process as projecting the set of signals onto a suitably chosen lower dimensional representation space, such that we get a more compact representation for the original signals while at the same time without losing information.

To see this, let us consider a simple example illustrated in Figure 2, where the signal class $\mathcal{X} = \bigcup_{i=1}^{3} \mathcal{S}_i$ is the union of three one-dimensional (1-D) subspaces. In the case of Figure 2(a), we

---

[1]The proposed sampling framework needs to be generalized in order to incorporate streams of Diracs, which do not belong to a Hilbert space. See [13] for details.
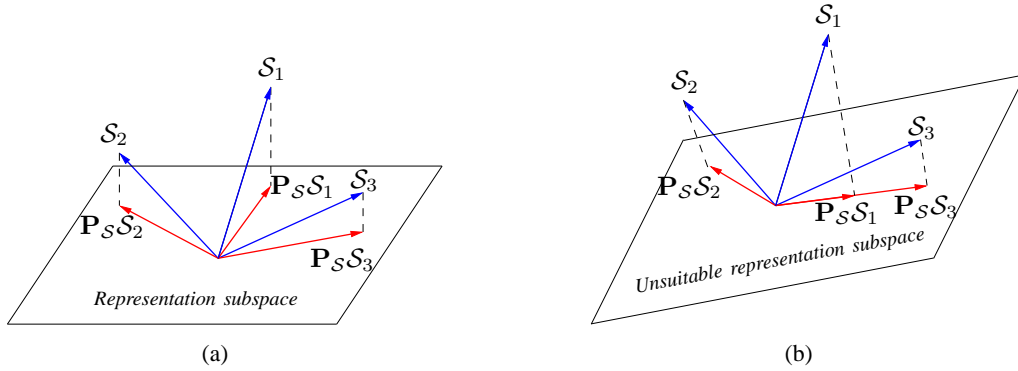
Fig. 2. Example of a union of subspaces $\mathcal{X} = \bigcup_{i=1}^{3} \mathcal{S}_i$ and its projections onto lower dimensional representation subspaces. (a) An invertible and stable representation. (b) A non-invertible representation. Also, a representation subspace close to this one would lead to unstable representation.

can see that there exists a one-to-one mapping between $\mathcal{X}$ and its projection $\mathbf{P}_{\mathcal{S}}\mathcal{X}$, since no two subspaces in $\{\mathcal{S}_i\}_{i=1}^{3}$ are projected onto a same line in $\mathcal{S}$. In contrast, Figure 2(b) shows a negative example, where the subspaces $\mathcal{S}_1$ and $\mathcal{S}_3$ are projected onto a single line in $\mathcal{S}$. Consequently, by merely looking at the projected (i.e. sampled) signal, we cannot tell which subspace ($\mathcal{S}_1$ or $\mathcal{S}_3$) it originally comes from.

The lower bound on the dimension of invertible representation subspaces is related to the minimum sampling requirement. In the case of Figure 2, we can see that the lowest possible dimension is two (i.e. a plane), because there would always be information loss if we project $\mathcal{X}$ onto any 1-D subspaces (i.e. single lines).

We can also see that the representation subspaces that provide invertible or one-to-one mapping are not unique. Although in theory any of them can be used, they are very different in practice. In particular, for those subspaces obtained through a small perturbation of Figure 2(b), the projected lines, albeit disjoint, would be so close to each other that sampling becomes very sensitive to noise and numerical error. So there is an issue in how to choose the "optimal" representation subspace, or equivalently the "optimal" sampling vectors.

III. CONDITIONS ON INVERTIBLE AND STABLE SAMPLING OPERATORS

Here we will present some main results regarding invertible and stable sampling operators. The proofs as well as various generalizations can be found in [13].

A. Two Sampling Requirements: Invertible and Stable Sampling

Recall the general sampling framework defined in Section II, where the set of signals of interest $\mathcal{X}$ is given in (3) and the sampling operator A is given in (4). First, we want to know whether each signal $x \in \mathcal{X}$ is uniquely represented by its sampling data Ax.

*Definition 1 (**Invertible sampling**):* We call A an *invertible* sampling operator for $\mathcal{X}$ if each $x \in \mathcal{X}$ is uniquely determined by its sampling data Ax; that means for every x_1 and x_2 in $\mathcal{X}$,

$$Ax_1 = Ax_2 \quad \text{implies} \quad x_1 = x_2.$$

In other words, A is a *one-to-one mapping* between $\mathcal{X}$ and $A\mathcal{X}$.

The invertible (or one-to-one) condition allows us to uniquely identify each $x \in \mathcal{X}$ from Ax. However, in practice, we want to be able to reconstruct $x \in \mathcal{X}$ in a *numerically stable* way from Ax. To guarantee such an algorithm exists, we need to ensure that if Ax is "close" to Ay then x is "close" to y as well [7]. Furthermore, we want that a small change in the signal x only produces a small change in its sampling data Ax. Consequently, we consider the following stricter condition on the sampling operator.

*Definition 2 (**Stable sampling**):* We call A a *stable* sampling operator for $\mathcal{X}$ if there exist two constants $\alpha > 0$ and $\beta < \infty$ such that for every $x_1 \in \mathcal{X}, x_2 \in \mathcal{X}$

$$\alpha \|x_1 - x_2\|_{\mathcal{H}}^2 \leq \|Ax_1 - Ax_2\|_{l_2(\Lambda)}^2 = \sum_{n \in \Lambda} |\langle x_1 - x_2, \psi_n \rangle|^2 \leq \beta \|x_1 - x_2\|_{\mathcal{H}}^2. \tag{5}$$

We call α and β *stability bounds* and the tightest ratio $\kappa = \beta/\alpha$ provides a measure for the stability of the sampling operator.

The main motivation behind the above stable sampling condition comes from the *frame theory* (see, for example, [14, pp. 53-63]). In fact, if the signal class $\mathcal{X}$ is a *single* vector space, then (5) becomes exactly the frame condition for the set of sampling vectors $\{\psi_n\}_{n \in \Lambda}$, with α and β being the lower and upper frame bounds, respectively. However, the difference here is that $\mathcal{X}$ forms a union of subspaces; therefore, x_1 and x_2 in (5) can be from two different subspaces.

The stable sampling condition is also related to the important concept of *restricted isometry property* (RIP), which is a key idea underlying the theory of compressive sampling [11], [12]. For example, when $\mathcal{X}$ represents finite-dimensional K-sparse signals, the inequalities in (5) are similar to the RIP condition on the measurement matrix whose row vectors are formed by the

sampling functions. The stable sampling condition (5) differs from the RIP in that we allow for an infinite number of subspaces in $\mathcal{X}$ (correspondingly, the measurement "matrix" will then have an infinite number of columns). Meanwhile, the signals we consider can be continuous-domain functions rather than discrete vectors.

B. Key Observation

The main difficulty in dealing with unions of subspaces is that the proposed unique and stable sampling conditions are defined on a *nonlinear* set. Consequently, we cannot directly apply various well-known *linear* results in matrix and operator theories to study the proposed sampling conditions. To overcome this problem, we introduce the following subspaces:

$$\widetilde{\mathcal{S}}_{\gamma,\theta} \stackrel{\text{def}}{=} \mathcal{S}_\gamma + \mathcal{S}_\theta = \{y : y = x_1 + x_2, \text{ where } x_1 \in \mathcal{S}_\gamma, \ x_2 \in \mathcal{S}_\theta\}, \tag{6}$$

It is easy to see that the set

$$\widetilde{\mathcal{X}} \stackrel{\text{def}}{=} \bigcup_{(\gamma,\theta) \in \Gamma \times \Gamma} \widetilde{\mathcal{S}}_{\gamma,\theta} = \{y : y = x_1 - x_2, \text{ where } x_1 \in \mathcal{X}, x_2 \in \mathcal{X}\}, \tag{7}$$

consists of all *secant* vectors of the set $\mathcal{X}$, which play a fundamental role in the study of dimensionality reduction [15].

By considering the new subspaces $\widetilde{\mathcal{S}}_{\gamma,\theta}$ in (6) and the relationship in (7), we can convert the two nonlinear conditions for invertible and stable sampling on a *union* of subspaces to a collection of linear conditions for *single* subspaces. In particular, we can show the following result [13]:

Proposition 1: A linear sampling operator A is *invertible* for $\mathcal{X}$ *if and only if* A is invertible for every $\widetilde{\mathcal{S}}_{\gamma,\theta}$, with $(\gamma, \theta) \in \Gamma \times \Gamma$. Furthermore, a linear sampling operator A is *stable* for $\mathcal{X}$ with stability bounds α and β, *if and only if*

$$\alpha\|y\|_{\mathcal{H}}^2 \le \|Ay\|_{l_2(\Lambda)}^2 \le \beta\|y\|_{\mathcal{H}}^2,$$

for every $y \in \widetilde{\mathcal{S}}_{\gamma,\theta}$ and $(\gamma, \theta) \in \Gamma \times \Gamma$.

C. Stability Bounds and Minimum Sampling Requirement

Recall that our linear sampling operator A is specified by a set of sampling vectors $\Psi = \{\psi_n\}_{n \in \Lambda}$ as defined in (4). This allows us to rephrase the invertible and stable sampling conditions

in terms of Ψ. For simplicity of exposition, we consider in the following the case where the subspaces $\mathcal{S}_\gamma$ $(\gamma \in \Gamma)$ in $\mathcal{X}$ are finite-dimensional. The extensions to cases where $\mathcal{S}_\gamma$ are infinite-dimensional shift-invariant subspaces can be found in [13].

Let $\Phi(\gamma, \theta) = \left\{ \phi_k^{(\gamma,\theta)} \right\}_{k=1}^{K(\gamma,\theta)}$ be an orthonormal basis for $\widetilde{\mathcal{S}}_{\gamma,\theta}$ with $\gamma, \theta \in \Gamma$, and let $\Psi = \{\psi_n\}_{n=1}^{N}$ represent a fixed set of N sampling vectors. The (generalized) Gram matrix $\boldsymbol{G}_{\Phi(\gamma,\theta),\Psi}$ between $\Phi(\gamma, \theta)$ and Ψ is defined as

$$\boldsymbol{G}_{\Phi(\gamma,\theta),\Psi} \stackrel{\mathrm{def}}{=} \left(\langle \phi_j^{(\gamma,\theta)}, \psi_i \rangle \right)_{\substack{i=1,\ldots,N \\ j=1,\ldots,K(\gamma,\theta)}} .$$

The following results [13] translate the invertible and stable sampling conditions on the sampling operator A into the (left) invertible condition on Gram matrices $\boldsymbol{G}_{\Phi(\gamma,\theta),\Psi}$.

*Proposition 2 (**Stability bounds**):* The sampling operator A provides an invertible sampling operator for $\mathcal{X}$ *if and only if* $\boldsymbol{G}_{\Phi(\gamma,\theta),\Psi}$ has full column rank for every $(\gamma, \theta) \in \Gamma \times \Gamma$. Furthermore, A provides a stable sampling operator for $\mathcal{X}$ *if and only if*

$$\alpha \stackrel{\mathrm{def}}{=} \inf_{(\gamma,\theta) \in \Gamma \times \Gamma} \sigma_{\min}^2(\boldsymbol{G}_{\Phi(\gamma,\theta),\Psi}) > 0, \quad \text{and}$$

$$\beta \stackrel{\mathrm{def}}{=} \sup_{(\gamma,\theta) \in \Gamma \times \Gamma} \sigma_{\max}^2(\boldsymbol{G}_{\Phi(\gamma,\theta),\Psi}) < \infty,$$

where $\sigma_{\min}(\boldsymbol{G})$ and $\sigma_{\max}(\boldsymbol{G})$ are the smallest and largest singular values of $\boldsymbol{G}$, respectively. Meanwhile, α and β are the tightest stability bounds.

*Proposition 3 (**Minimum sampling requirement**):* Suppose that $A : x \mapsto \{\langle x, \psi_n \rangle\}_{n=1}^{N}$ is an *invertible* sampling operator for $\mathcal{X}$. Then

$$N \geq N_{\min} \stackrel{\mathrm{def}}{=} \sup_{(\gamma,\theta) \in \Gamma \times \Gamma} \dim(\widetilde{\mathcal{S}}_{\gamma,\theta}). \tag{8}$$

Note that, when the index set Γ is countable, the above lower bound on the number of required samples can be achieved, as is shown by the following result [13].

Proposition 4 (Existence of Invertible Sampling Operators): Suppose that $\mathcal{X} = \bigcup_{\gamma \in \Gamma} \mathcal{S}_\gamma$ is a *countable* union of subspaces of $\mathcal{H}$, and that $N_{\min}$ in (8) is finite. Then the collection of sampling vectors $\left\{ \Psi = \{\psi_n\}_{n=1}^{N_{\min}} : \Psi \text{ forms an invertible sampling operator for } \mathcal{X} \right\}$ is dense in $\mathcal{H}^{N_{\min}}$.

For the special case of finite-dimensional K-sparse signals, both the minimum sampling requirement in Proposition 3 and the existence argument in Proposition 4 reduce to established

results in the compressive sampling literature. The generalization of the above propositions lies in the fact that we allow the ambient space $\mathcal{H}$ to be infinite-dimensional and the index set Γ to also be infinite; also, the result of Proposition 3 can be extended [13] to cases when the signals are continuous-domain functions living in a union of infinite-dimensional shift-invariant subspaces.

IV. APPLICATIONS

We finish our discussions by describing several applications of the theoretical results stated in the last section.

First, consider $\mathcal{X}$ to be a class of signals supported on $[0, 1]$, where each signal consists of K different polynomials pieces of degree less than d (see Figure 3 for an example). Clearly, every piecewise polynomial signal in $\mathcal{X}$ can be fully specified by $Kd + K - 1$ parameters, with $K - 1$ parameters used to record the locations of the discontinuities and Kd parameters to specify the K polynomial pieces. But is it sufficient to use only $Kd + K - 1$ linear measurements to fully specify signals from $\mathcal{X}$?

This question can be answered by applying Proposition 3. We can check that $\widetilde{\mathcal{S}}_{\gamma,\theta}$ as defined in (6) are subspaces of piecewise polynomials with at most $2K - 1$ pieces, each of degree less than d. It follows from (8) that the minimum sampling requirement for $\mathcal{X}$ is $N_{\min} = (2K - 1)d$, which is strictly greater than the number of free parameters $Kd + K - 1$ when $d > 1$. Thus, as a novel application of our minimum sampling bound, we have shown that the sampling algorithm proposed in [5] for piecewise polynomials has indeed achieved the minimum sampling requirement $N_{\min} = (2K - 1)d$.

Next, let us revisit Example 3 of sparse representations. It is well-known in the compressed sensing/compressive sampling literature that finite-dimensional K-sparse signals can be uniquely determined by a minimum of $2K$ linear measurements. As an application of Proposition 4, we extend the above standard result to the infinite-dimensional case. Suppose $\mathcal{H}$ is a separable Hilbert space (e.g. $L^2(\mathbb{R})$) and let $\{\phi_k\}_{k=1}^{\infty}$ be a countable basis for $\mathcal{H}$. Then the set $\mathcal{X}$ of all possible K-term representations as given in (2) constitutes a countable union of subspaces of dimension K in $\mathcal{H}$. Proposition 4 guarantees that the collection of $2K$ vector sets $\{\psi_n\}_{n=1}^{2K}$ that provide invertible sampling operators not only exists, but also forms a *dense* set.

Finally, we mention a concrete example of stable sampling vectors. Recall the sampling problem described in Example 2, where the signals class $\mathcal{X}^{MB}$ is a union of infinite-dimensional
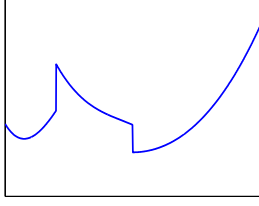
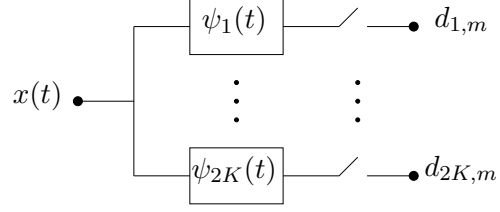Fig. 3. A piecewise polynomial with unknown discontinuity locations.

Fig. 4. Multi-channel sampling. The input signal first goes through a bank of $2K$ filters, and then the samples are taken at time instances $\{mT_0 : m \in \mathbb{Z}\}$.

bandlimited subspaces. Using a result similar to Proposition 3, we can show that the minimum sampling rate here is $2K$ samples per unit time $T_0 \overset{\text{def}}{=} L/(f_{\max} - f_{\min})$, which is twice the rate we would need if we possessed prior knowledge about the frequency support. Interestingly, this minimum rate can be achieved by a large class of sampling vectors providing stable sampling. In particular, consider the sampling procedure A shown in Figure 4, where signals from $\mathcal{X}^{MB}$ go through a bank of $2K$ filters $\{\psi_n(t)\}_{n=1}^{2K}$ followed by uniform sampling. We show in [13] that a sufficient condition for A to be a stable sampling operator as defined in (5) is that the Fourier transforms of the sampling functions $\left\{\widehat{\psi}_n(f)\right\}_{n=1}^{2K}$ are continuous on the interval $[f_{\min}, f_{\max}]$ and satisfy

$$\det \begin{pmatrix} \widehat{\psi}_1(f_1) & \widehat{\psi}_1(f_2) & \cdots & \widehat{\psi}_1(f_{2K}) \\ \widehat{\psi}_2(f_1) & \widehat{\psi}_2(f_2) & \cdots & \widehat{\psi}_2(f_{2K}) \\ \vdots & \vdots & \ddots & \vdots \\ \widehat{\psi}_{2K}(f_1) & \widehat{\psi}_{2K}(f_2) & \cdots & \widehat{\psi}_{2K}(f_{2K}) \end{pmatrix} \neq 0, \tag{9}$$

for all choices of $0 \leq f_{\min} \leq f_1 < f_2 < \ldots < f_{2K} < f_{\max}$.

There are many different sampling functions for which (9) holds. For example, based on the property of the Vandermonde matrices, we can verify [13] that $\widehat{\psi}_n(f) = e^{j\frac{2\pi(n-1)}{f_{\max}-f_{\min}}f}$ $(n = 1\ldots 2K)$ (which corresponds to the original periodic nonuniform sampling scheme in [8]) and $\widehat{\psi}_n(f) = e^{-f^2 n/\sigma^2}$ $(n = 1\ldots 2K)$ are just two special choices.

V. CLOSING REMARKS

In this article, we provided a new perspective for signal sampling, by considering signals from a union of subspaces instead of a single space. We have shown the conditions for invertible and stable sampling; and derived the minimum sampling requirement within this framework, which is typically more efficient than the minimum sampling rate dictated by classical sampling results

considering only single spaces.

The proposed union of subspaces model is fairly general, and encompasses many basic signal classes studied in the related work on spectrum-blind sampling, finite rate of innovation sampling, and compressed sensing/compressive sampling. It is our hope that the proposed sampling framework can serve as a common ground, on which techniques and lessons can be learned from the above three lines of thinking, and applied to the sampling of other general signal classes (e.g. a union of shift-invariant subspaces) that have not been previously addressed in the literature.

REFERENCES

[1] C. E. Shannon, "Communications in the presence of noise," *Proc. IRE*, vol. 37, pp. 10–21, 1949.

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

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

[4] A. M. Bruchstein, T. J. Shan, and T. Kailath, "The resolution of overlapping echos," *IEEE Trans. Acoust., Speech, and Signal Process.*, vol. 33, no. 6, pp. 1357–1367, Dec. 1985.

[5] M. Vetterli, P. Marziliano, and T. Blu, "Sampling signals with finite rate of innovation," *IEEE Trans. Signal Process.*, vol. 50, no. 6, pp. 1417–1428, Jun. 2002.

[6] P. Dragotti, M. Vetterli, and T. Blu, "Sampling moments and reconstructing signals of finite rate of innovation: Shannon meets Strang-Fix," *IEEE Trans. Signal Process.*, vol. 55, no. 5, pp. 1741–1757, May 2007.

[7] H. J. Landau, "Sampling, data transmission, and the Nyquist rate," *Proc. IEEE*, vol. 55, no. 10, pp. 1701–1706, Oct. 1967.

[8] P. Feng and Y. Bresler, "Spectrum-blind minimum-rate sampling and reconstruction of multiband signals," in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Proc.*, vol. 3, Atlanta, GA, May 1996, pp. 1689–1692.

[9] Y. Bresler, M. Gastpar, and R. Venkataramani, "Image compression on-the-fly by universal sampling in Fourier imaging systems," in *IEEE Information Theory Workop on Detection, Estimation, Classification, and Imaging*, Santa-Fe, USA, 1999.

[10] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, pp. 1289–1306, Apr. 2006.

[11] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, pp. 489–509, Feb. 2006.

[12] E. J. Candès and T. Tao, "Near optimal signal recovery from random projections: Universal encoding strategies?" *IEEE Trans. Inf. Theory*, vol. 52, pp. 5406–5425, Dec. 2006.

[13] Y. M. Lu and M. N. Do, "A theory for sampling signals from a union of subspaces," *IEEE Trans. Signal Process.*, to appear.

[14] I. Daubechies, *Ten Lectures on Wavelets*. Philadelphia, PA: SIAM, 1992.

[15] D. S. Broomhead and M. Kirby, "A new approach to dimensionality reduction: Theory and algorithms," *SIAM J. Appl. Math.*, vol. 60, no. 6, pp. 2114–2142, 2000.