A GEOMETRICAL APPROACH TO SAMPLING SIGNALS WITH FINITE RATE OF INNOVATION

Yue Lu and Minh N. Do

Department of Electrical and Computer Engineering
Coordinated Science Laboratory
University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA
E-Mail: {yuelu, minhdo}@uiuc.edu

ABSTRACT

Many signals of interest can be characterized by a finite number of parameters per unit of time. Instead of spanning a single linear space, these signals often lie on a union of spaces. Under this setting, traditional sampling schemes are either inapplicable or very inefficient. We present a framework for sampling these signals based on an injective projection operator, which "flattens" the signals down to a common low dimensional representation space while still preserves all the information. Standard sampling procedures can then be applied on that space. We show the necessary and sufficient conditions for such operators to exist and provide the minimum sampling rate for the representation space, which indicates the efficiency of this framework. These results provide a new perspective on the sampling of signals with finite rate of innovation and can serve as a guideline for designing new algorithms for a class of problems in signal processing and communications.

1. INTRODUCTION

Many signals of interest can be characterized by a finite number of parameters (degrees of freedom) per unit of time. Following the convention in [1], we call them signals with finite rate of innovation. Instead of spanning a single linear space, these signals often lie on a union of spaces, as shown in the following examples.

Example 1 (Stream of Diracs) *Consider a stream of K Diracs periodized with period T. Within each period, the signal $x(t) = \sum_{k=1}^{K} c_k \delta(t - t_k)$, where $\{t_k\}_{k=1}^{K}$ can be chosen arbitrarily from $[0, T)$.*

we see once the K locations are fixed, the signals span a K dimensional subspace within one period. While with locations unknown, we have a union of spaces.

Example 2 (2D Piecewise Polynomials) *Now we consider the discrete periodic 2-D piecewise polynomials of K pieces, as shown in Figure 1. More specifically, each piece is a bi-variate polynomial of degree less than d.*

This kind of signals can be seen as a "cartoon" model of natural images, since natural scenes are often made up from several objects with smooth boundaries. Again, the signals lie on a union of subspaces corresponding to all possible boundaries.

Example 3 (Signals with Unknown Spectral Support) *[2] In this example, we consider the class of continuous-time signals bandlimited to the spectral span $\mathcal{F} = [f_{min}, f_{max}]$. We only know the spectral support of the signals occupies a certain portion (say $1/5$) of $\mathcal{F}$, but we do not know the exact location of the support.*

For a fixed frequency support, these signals become a (shift-invariant) subspace. When the support is unknown, the signal class can be characterized as a union of subspaces.

Before going to the formal definition for signals with finite rate of innovation, we will first look at the concept of shift-invariant spaces [3, 4], which delineate the scope of this paper.

Definition 1 (Shift-Invariant Spaces) *$\mathcal{S}$ is called a (finitely generated) shift-invariant space, if*

$$\mathcal{S} = \left\{ \sum_{n \in \mathbb{Z}} \sum_{r=1}^{D} c_{nr} \varphi_r\left(\frac{t}{T} - n\right) \right\} \quad (1)$$

where $\{\varphi_r\}_{r=1}^{D}$ are the generating functions of $\mathcal{S}$ and D is called the length of $\mathcal{S}$, denoted by $\text{len}(\mathcal{S})$.

Familiar examples of shift-invariant spaces include bandlimited signals, and uniform splines. A particularly interesting case is when the generating functions $\varphi_r(\frac{t}{T})$ are compactly supported within a certain period, say $[0, T]$. In this case, signal segments in different periods are independent of each other. While in each period, these signal segments belong to a finite dimensional space. Note that periodic signals can be seen as a special case under this setting, if we further impose $c_{nr} = c_{kr}, \forall n, k \in \mathbb{Z}, r = 1 \ldots d$.

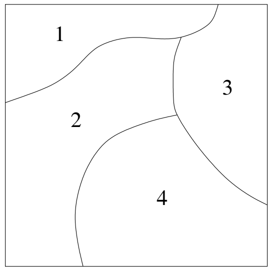

Fig. 1. 2D Piecewise Polynomials.

Definition 2 (Signals with Finite Rate of Innovation) [1] *Let $\mathcal{M}$ denote a class of signals in a Hilbert space $\mathcal{H}$. We call $\mathcal{M}$ having finite rate of innovation, if it can be written as a union of finitely generated shift-invariant subspaces. That means $\mathcal{M} = \bigcup_{\gamma \in \mathbf{\Gamma}} \mathcal{S}_\gamma$, where $\{\mathcal{S}_\gamma\}_{\gamma \in \mathbf{\Gamma}}$ is a collection of shift-invariant spaces with $\operatorname{len}(\mathcal{S}_\gamma) \leq D$ and the index set $\mathbf{\Gamma} \subseteq \mathbb{R}^m$.*

In this paper, we present a framework for sampling this class of signals. We formulate the problem in Section 2 with a geometrical perspective relating sampling to the projection operators onto representation spaces. The key observation here is that the original signals can be fully recovered from their sampled values if the projector satisfies certain conditions. In Section 3, we present the necessary and sufficient conditions for the projectors. As a direct result, we also show the minimum sampling rate. In [1], the authors propose the sampling scheme for certain special signals, such as the stream of Diracs and piecewise polynomials. Developed within a very general setting, the results in this paper indicate that actually a very broad class of signals with finite rate of innovation can be (uniformly) sampled without loss of information. Finally, we also (briefly) discuss how these results can serve as a guideline for designing new sampling algorithms with more stability and robustness against the noise.

2. SIGNAL REPRESENTATION THROUGH A PROJECTION OPERATOR

From Definition 2, we see once the index γ is known, the signals lie on a single space and traditional sampling scheme can be applied here. However, in practice it is often very difficult to get these index parameters. For example, it is not a trivial task to locate the discontinuities of a piecewise polynomial signal, especially in the discrete case and with noise. In this sense, we want to have a fixed sampling scheme that works for all signals in $\mathcal{M}$ without having any knowledge about which particular space the signals belong to.

[1]The difference between this definition and the one proposed in [1] is discussed in [5].

Let $x(t)$ denote the original signal. The traditional uniform sampling setup is to filter $x(t)$ with a certain sampling kernel (e.g. $sinc$), and get the samples $y(nT)$ by

$$y(nT) = \int_{-\infty}^{+\infty} h(t-nT)x(t)dt = \langle h(t-nT), x(t) \rangle, \quad (2)$$

where $\tilde{h}(t) = h(-t)$ is the sampling kernel and T is the time interval. We see the samples $y(nT)$ are actually the inner products between $x(t)$ and $h(t-nT)$. Let $\mathcal{S}$ be the shift invariant space generated by $h(tT)$, i.e. $\varphi = h(tT)$ in Definition 1. It can be shown [5] that the sampling process is equivalent (in the sense of one-to-one correspondence) to the projection of the signals onto $\mathcal{S}$. So in the following discussion, we can just focus on the problem of projection.

Of course, we want that the original signals can be perfectly reconstructed from their sampled values. If we think the sampling process as a mapping from the original signal to the sample values, then it must be an invertible mapping. We characterize this constraint through the following definition.

Definition 3 *Let $\mathcal{M}$ denote a class of signals in the signal space $\mathcal{H}$. We call an orthogonal projection operator $\mathbf{P}_\mathcal{S} : \mathcal{H} \longmapsto \mathcal{S}$ a representation for $\mathcal{M}$, if any signal $\boldsymbol{m} \in \mathcal{M}$ can be uniquely determined by its image $\mathbf{P}_\mathcal{S}\boldsymbol{m}$, i.e. .*

$$\mathbf{P}_\mathcal{S}\boldsymbol{m}_1 = \mathbf{P}_\mathcal{S}\boldsymbol{m}_2 \quad \textit{if and only if} \quad \boldsymbol{m}_1 = \boldsymbol{m}_2 \quad (3)$$

Now the natural questions to pursue are the following.

1. What are the necessary and sufficient conditions for such projection operators to exist?

2. What is the efficiency of such a representation? Or equivalently, what is the minimum sampling rate for this class of signals?

Before going to the formal results in the next section, let us fist consider a very simple case, from which we can gain some valuable geometrical intuition. Here the signal space $\mathcal{H}$ is $\mathbb{R}^3$. $\mathcal{S}_1$, $\mathcal{S}_2$, and $\mathcal{S}_3$ are the one-dimensional subspaces corresponding to three disjoint lines, and $\mathcal{M} = \cup_{i=1}^3 \mathcal{S}_i$. As shown in Fig. 2, we project $\mathcal{M}$ down to a certain space and get $\mathbf{P}\mathcal{M} = \cup_{i=1}^3 \mathbf{P}\mathcal{S}_i$, where $\mathbf{P}$ is the projection operator. We can see that there is a one-to-one mapping between $\mathcal{M}$ and $\mathbf{P}\mathcal{M}$, as long as no two subspaces are projected onto a same line. In this case, no information is lost while we have a more compact representation of the original signals. Intuitively, we can think of this process as "flattening" down the signals to a low dimensional representation space, while still preserving the original structural information. It is interesting to study the lower bound of the dimension of the representation space, since the lower the dimension, the more efficient the representation is. The lower bound is 2

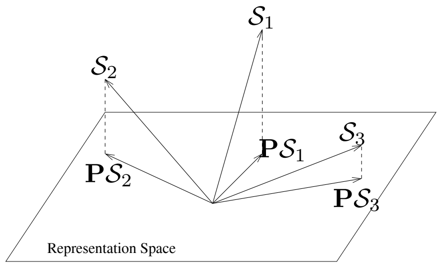

Fig. 2. The projection operator "flattens" the original signals down to a low dimensional representation space, while the structural information is still preserved.

(i.e. a plane) in this case, because there must be information loss if we project $\mathcal{M}$ onto any single line.

It is also straightforward to notice that the projection operators that ensure one-to-one mapping are not unique. Though in principle any of them can be used, they are very different in the practical sense. For some projectors, the projected lines are so close to each other that the system becomes very sensitive to noise and numerical imprecision. So there is an issue in how to choose the "optimal" projection operator. We will formalize the above geometric intuition in the next section.

3. THE NECESSARY AND SUFFICIENT CONDITIONS FOR THE PROJECTION (SAMPLING) OPERATORS

Section 2 relates the sampling of signals with finite rate of innovation to the problem of projecting signals onto representation spaces. In this section, we will study the conditions those projection operators must satisfy to ensure there is no information loss in the process. For simplicity reasons, we only discuss the case of periodic signals, where the signals lie on a union of finite-dimensional spaces in one period. Actually these results can be easily extended to the case of general shift-invariant spaces [5].

3.1. Injective Projection Operators

Theorem 1 *M is a class of signals in $\mathcal{H}$ having finite rate of innovation. A projection operator $\mathbf{P}$ satisfies* (3) *if and only if for all $\widetilde{\mathcal{S}}_{\boldsymbol{ij}} \triangleq \mathcal{S}_{\boldsymbol{i}} + \mathcal{S}_{\boldsymbol{j}}$ with $(\boldsymbol{i}, \boldsymbol{j}) \in \boldsymbol{\Gamma} \times \boldsymbol{\Gamma}$, $\mathbf{P}$ is an injective operator (one-to-one mapping) from $\widetilde{\mathcal{S}}_{\boldsymbol{ij}}$ onto $\mathbf{P}\widetilde{\mathcal{S}}_{\boldsymbol{ij}}$.*

Remark 1 *Here $\mathbf{P}\widetilde{\mathcal{S}}_{\boldsymbol{ij}}$ denotes the image of $\widetilde{\mathcal{S}}_{\boldsymbol{ij}}$ under $\mathbf{P}$. Since $\mathbf{P}$ is a linear operator, $\mathbf{P}\widetilde{\mathcal{S}}_{\boldsymbol{ij}}$ is still a subspace. We define $\widetilde{\mathcal{S}}_{\boldsymbol{ij}}$ as the sum, instead of the direct sum, of $\mathcal{S}_{\boldsymbol{i}}$ and $\mathcal{S}_{\boldsymbol{j}}$, since we do not require the two subspaces to be disjoint.*

Proof 1 *The sufficiency of the condition is easy to show. For any pair $\boldsymbol{m}_1, \boldsymbol{m}_2 \in \mathcal{M}$ such that $\mathbf{P}\boldsymbol{m}_1 = \mathbf{P}\boldsymbol{m}_2$. Without loss of generality, suppose $\boldsymbol{m}_1 \in \mathcal{S}_1$, $\boldsymbol{m}_2 \in \mathcal{S}_2$ and hence $\boldsymbol{m}_1 \in \widetilde{\mathcal{S}}_{12}, \boldsymbol{m}_2 \in \widetilde{\mathcal{S}}_{12}$. Since $\mathbf{P}$ is an injective operator from $\widetilde{\mathcal{S}}_{12}$ to $\mathbf{P}(\widetilde{\mathcal{S}}_{12})$, we must have*

$$\mathbf{P}\boldsymbol{m}_1 = \mathbf{P}\boldsymbol{m}_2 \Rightarrow \boldsymbol{m}_1 = \boldsymbol{m}_2 \qquad (4)$$

Next, we will show the condition is also necessary by contradiction. Suppose $\mathbf{P}$ satisfies (3)*, but there exists some $\widetilde{\mathcal{S}}_{12} = \mathcal{S}_1 + \mathcal{S}_2$ such that $\mathbf{P}$ is not injective from $\widetilde{\mathcal{S}}_{12}$ to $\mathbf{P}(\widetilde{\mathcal{S}}_{12})$. Let $\{\varphi_k\}_{k=1}^{K}$ denote a certain basis for $\widetilde{\mathcal{S}}_{1,2}$ such that $\{\varphi_k\}_{k=1}^{n_0} \in \mathcal{S}_1$ and $\{\varphi_k\}_{k=n_0+1}^{K} \in \mathcal{S}_2$ for some n_0. Note that such a basis can always be found. Since $\mathbf{P}$ is not injective, there must be some $\boldsymbol{s}_1 = \sum_{k=1}^{K} \alpha_k \varphi_k$ and $\boldsymbol{s}_2 = \sum_{k=1}^{K} \beta_k \varphi_k$ with $\boldsymbol{s}_1 \neq \boldsymbol{s}_2$ but $\mathbf{P}\boldsymbol{s}_1 = \mathbf{P}\boldsymbol{s}_2$. Now we have*

$$\mathbf{P}\left(\sum_{k=1}^{K} \alpha_k \varphi_k\right) = \mathbf{P}\left(\sum_{k=1}^{K} \beta_k \varphi_k\right) \qquad (5)$$

It then follows from the linearity of $\mathbf{P}$ that

$$\mathbf{P}\left(\sum_{k=1}^{n_0} (\alpha_k - \beta_k)\varphi_k\right) = \mathbf{P}\left(\sum_{k=n_0+1}^{K} (\beta_k - \alpha_k)\varphi_k\right) \quad (6)$$

Now let $\boldsymbol{m}_1 = \sum_{k=1}^{n_0}(\alpha_k - \beta_k)\varphi_k$ and $\boldsymbol{m}_2 = \sum_{k=n_0+1}^{K}(\beta_k - \alpha_k)\varphi_k$. It is clear that $\boldsymbol{m}_1 \neq \boldsymbol{m}_2$, and $\boldsymbol{m}_1 \in \mathcal{S}_1, \boldsymbol{m}_2 \in \mathcal{S}_2$. From (6)*, $\mathbf{P}\boldsymbol{m}_1 = \mathbf{P}\boldsymbol{m}_2$. This contradicts the assumption that $\mathbf{P}$ satisfies* (3)*. Therefore we must have $\mathbf{P}$ to be injective from $\widetilde{\mathcal{S}}_{\boldsymbol{i},\boldsymbol{j}}$ to $\mathbf{P}(\widetilde{\mathcal{S}}_{\boldsymbol{i},\boldsymbol{j}}), \forall(\boldsymbol{i}, \boldsymbol{j}) \in \boldsymbol{\Gamma} \times \boldsymbol{\Gamma}$.* ■

Corollary 1 (Minimum Sampling Rate) *$\mathcal{M}$ is a class of signals with finite rate of innovation. Let $D_M = \max\limits_{(\boldsymbol{i},\boldsymbol{j}) \in \boldsymbol{\Gamma} \times \boldsymbol{\Gamma}} D_{\boldsymbol{ij}}$, where $D_{\boldsymbol{ij}} \triangleq \dim(\widetilde{\mathcal{S}}_{\boldsymbol{ij}})$. A necessary condition for $\mathbf{P} : \mathcal{H} \longmapsto \mathcal{S}$ to be a representation for M is $\dim(\mathcal{S}) \geq D_M$.*

Remark 2 *Basically $\mathbf{P} : \mathcal{H} \longmapsto \mathcal{S}$ allows us to use $D = \dim(\mathcal{S})$ real numbers to represent the signals in $\mathcal{M}$. Of course the smaller the number D, the more efficient the representation can be. What Corollary 1 tells us is that the lower bound of D is D_M. It is impossible to use less than D_M numbers to fully represent the signal class $\mathcal{M}$.*

As an application of the Corollary, let us revisit the example of piecewise polynomial signals. For simplicity of exposition, we just consider the 1-D case. Within one period, the signals contain K pieces of polynomials, each of degree less than d. Intuitively, we should be able to represent any such signal by only $Kd + K - 1$ numbers, in which $K - 1$ numbers are used to locate the discontinuities, and Kd to specify the polynomial coefficients. However, we can verify that $D_M = (2K - 1)d$ in this case. It follows from Corollary 1 that the minimum number of samples we need is $(2K - 1)d$, which is always strictly larger than $Kd + K - 1$ when $d > 1$.

3.2. An Equivalent Condition

The following equivalent condition is usually more useful in practice than the one in Theorem 1.

Proposition 1 *M is a class of signals in $\mathcal{H}$ having finite rate of innovation. A projection operator $\mathbf{P} : \mathcal{H} \longmapsto \mathcal{S}$ satisfies* (3) *if and only if for all $\widetilde{\mathcal{S}}_{\boldsymbol{ij}}$ with $(\boldsymbol{i}, \boldsymbol{j}) \in \boldsymbol{\Gamma} \times \boldsymbol{\Gamma}$, the following matrix*

$$\boldsymbol{G} = \begin{pmatrix} \langle\varphi_1, \phi_1\rangle & \langle\varphi_2, \phi_1\rangle & \cdots & \langle\varphi_N, \phi_1\rangle \\ \langle\varphi_1, \phi_2\rangle & \langle\varphi_2, \phi_2\rangle & \cdots & \langle\varphi_N, \phi_2\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle\varphi_1, \phi_D\rangle & \langle\varphi_2, \phi_D\rangle & \cdots & \langle\varphi_N, \phi_D\rangle \end{pmatrix} \quad (7)$$

has full column rank, where $\{\varphi_i\}_{i=1}^N$ are the basis vectors of $\widetilde{\mathcal{S}}_{\boldsymbol{ij}}$ and $\{\phi_i\}_{i=1}^D$ are the basis vectors of $\mathcal{S}$.

We can see that a necessary condition for any $D \times N$ matrix $\boldsymbol{G}$ to have full column rank is $D \geq N$. This leads to the same result in Corollary 1 for the minimum dimension (sampling rate) of the projection space.

To show how to make use of the above results, let us reexamine the periodic signals containing stream of Diracs in Example 1. In [1], Vetterli etc. propose a sampling scheme for these signals by using the $sinc$ kernel. It can be seen as projecting the signals onto the space spanned by D basis vectors $\{e^{j(2\pi kt)/T}\}_{k=1}^D$. Actually, we will see this is just one of the many possible solutions.

First, the basis vectors for $\widetilde{\mathcal{S}}_{\boldsymbol{ij}}$ are $\{\varphi_i = \delta(t - t_i)\}_{i=1}^D$, with $0 \leq t_1 < t_2 < \ldots < t_D < T$ and $D = 2K$. If we specifically choose the basis vectors of the projection space to be $\phi_1 = f(t)$,$\phi_2 = f^2(t), \ldots, \phi_D = f^D(t)$ with $f(t)$ an arbitrary function defined on $[0, T]$, then the inner product $\langle\varphi_i, \phi_j\rangle = \langle\delta(t - t_i), f^j(t)\rangle = f^j(t_i)$. The matrix $\boldsymbol{G}$ can now be written as

$$\boldsymbol{G} \;=\; \begin{pmatrix} f(t_1) & f(t_2) & \cdots & f(t_D) \\ f^2(t_1) & f^2(t_2) & \cdots & f^2(t_D) \\ \vdots & \vdots & \ddots & \vdots \\ f^D(t_1) & f^D(t_2) & \cdots & f^D(t_D) \end{pmatrix}, \quad (8)$$

which is a Vandermonde system. To ensure $\boldsymbol{G}$ nonsingular (and hence having full column rank), the only requirement is $f(t_i) \neq f(t_j)$, for all $0 \leq t_i < t_j < T$ [6]. It is easy to verify that $f(t) = e^{j(2\pi t)/T}$ is one of the functions satisfying the above requirement. Since there are lots of other choices of suitable $f(t)$, it is possible to find some $f(t)$ having better performance, e.g. with shorter support or more robustness against the noise.

3.3. The Existence of Injective Projections

Using the classical *Baire's Theorem* [7] in analysis, we can show that the minimum sampling rate can always be achieved when the index set $\boldsymbol{\Gamma}$ is countable, as stated by the following theorem.

Theorem 2 *Let a Hilbert space $\mathcal{H}$ denote the signal space. Let $\mathcal{M}$ be an arbitrary class of signals in $\mathcal{H}$ having finite rate of innovation with a countable index set $\boldsymbol{\Gamma}$. We can always find a representation for $\mathcal{M}$ through a projection operator $\mathbf{P}$ from $\mathcal{H}$ to $\mathcal{S}$ satisfying* (3)*, where $\dim(\mathcal{S}) = \max_{(\boldsymbol{i},\boldsymbol{j})\in\boldsymbol{\Gamma}\times\boldsymbol{\Gamma}} \dim(\widetilde{\mathcal{S}}_{\boldsymbol{ij}})$ is the minimum sampling rate. Furthermore, those feasible operators form a dense set in the space of linear operators.*

4. CONCLUSION AND FUTURE WORK

We considered the sampling of signals with finite rate of innovation. The key idea is to model those signals as a union of shift-invariant spaces and find a suitable projection operator which "flattens" the signals down to a low dimensional representation space while still preserves all the information. We discussed the necessary and sufficient conditions for such operators to exist. Meanwhile, we also provided the minimum sampling rate of these signals. The results in the paper are developed within a very general setting and provide a new perspective on the problem of sampling signals with finite rate of innovation. The insight and various conditions developed here can serve as a guideline for designing new algorithms for a class of related problems in signal processing and communications.

5. REFERENCES

[1] M. Vetterli, P. Marziliano, and T. Blu, "Sampling signals with finite rate of innovation," *IEEE Trans. Signal Proc.*, vol. 50, pp. 1417–1428.

[2] 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.*, Atlanta, USA, 1996.

[3] M. Unser, "Sampling — 50 years after shannon," *Proc. IEEE*, vol. 88, no. 4, pp. 569–587, April 2000.

[4] C. de Boor, R. A. DeVore, and A. Ron, "The structure of finitely generated shift-invariant spaces in $L_2(\mathbb{R}^d)$," *J. of Functional Analysis*, vol. 119, pp. 37–78, 1994.

[5] Y. Lu and M. N. Do, "Representing signals with finite rate of innovation: An injective projection approach," Tech. Rep.

[6] R. Horn and C. Johnson, *Matrix Analysis*, Cambridge University Press, 1985.

[7] H. L. Royden, *Real Analysis*, Prentice Hall, 1988.