

On Minimum Sampling Rates for Signals in Shift Invariant Spaces

Beilei Huang, Edmund M-K. Lai
 Institute of Information Sciences & Technology
 Massey University, Wellington, New Zealand

A.P. Vinod
 School of Computer Engineering
 Nanyang Technological University, Singapore

Abstract—This paper is focus on to design a sampling system with minimum sampling rate for signals in the shift invariant Hilbert space. To achieve this goal, we propose a method to calculate the Rate of Innovation (RI) of signals in the Hilbert space. The RI is then used to identify a suitable sampling kernel as well as the sampling rate for a specific signal. We show that the RI of the kernel should be greater or equal to the RI of the signal for the signal to be perfectly reconstructible. The minimum sampling rate depends on the RI of the signal. Examples are included to demonstrate how our method is applied to calculate the RI. Some known sampling theories can also be fitted into the framework of sampling system with minimum sampling rate.

I. INTRODUCTION

Sampling and reconstruction are the most fundamental techniques in signal processing applications. A continuous signal is represented by its discrete samples and further processing of the signal can be carried out in the digital domain. To sample a signal efficiently implies to represent the signal using minimum number of samples, yet the signal can be perfectly reconstructed afterwards. For bandlimited signals $f(x)$, where $F(\omega) = 0$ for $\omega > \omega_0$, the minimum number of samples are obtained by using the kernel $h(x) = \text{sinc}Bt$ where $B = 2\omega_0/2\pi$ and sampling at $T = 1/B$. For non-bandlimited signal, unfortunately, the answer is not straight available.

A benchmark result from [1] states that the signal can be sampled at their Rate of Innovation (RI) using an appropriate kernel and then be perfectly reconstructed, despite their non-bandlimited response. The RI is defined as number of unknowns per unit time. Consider the periodic stream of Diracs for example: $f(x) = \sum_n c_n \delta(x - x_n)$ where $c_n = c_{n+K}$ and $x_n = x_{n+K} - \tau$, τ is the periodicity and K is number of pulses per cycle; the RI of $x(t)$ is $\rho = 2K/\tau$. The key of sampling resides in to identify the innovative part of the signal, like the time locations and the weights of the pulses, and to reconstruct accordingly. The essentialness of RI is that it indicates the minimum sampling rate associated with a certain sampling kernel. The value of ρ may be intuitive for some signals, however, it is not so straightforward for other non trivial signals.

In this paper we propose a method to calculate the RI of a general signal in shift invariant Hilbert space. Based on that, we design a sampling system with suitable sampling kernel and minimum sampling rate for that specific signal. In this manner, the signal is described with minimum number of samples, yet perfectly reconstructible. As we illustrate in

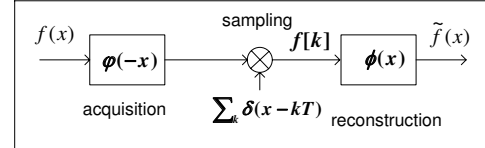


Fig. 1. Block diagram representation sampling system. The meaning of symbols can be found in Section II.

Section IV, the choice of the sampling kernel depends on the relationship of the RI of the signal's and the sampling kernel's. We show through experimental results that our proposed method to calculate the RI is accurate; and the way to identify a suitable sampling kernel to have minimum sampling rate for a signal is valid. Through the paper the signals and all sampling kernels are restricted to be in the shift invariant Hilbert Space. We use ρ_f to denote the RI of the signal $f(x)$.

II. SAMPLING SYSTEM REVIEWED

Sampling theory is the most fundamental topic in signal processing. Ever since Shannon's sampling theory on bandlimited signals, the sampling theory nowadays has been extended to deal with multiband, non-bandlimited and non-uniform sampled signals as well [2]–[4]. Recently, thanks to the connection to frame theory, the splines and wavelet are also intensively applied in sampling systems [5], [6]. The sampling process is reformed to express the signal in a shift invariant space spanned by the frame. The schematic diagram of a general sampling system is shown in Figure 1.

Assume that the signal $f(x)$ is in Hilbert square summable space and $\varphi(-x)$ and $\phi(x)$ are the acquisition and reconstruction filter respectively. In the acquisition and sampling stage, the signal is passed through $\varphi(-x)$ and sampled at a

$$f[k] = \langle f(x), T_{ka}\varphi(x) \rangle \quad (1)$$

T defines the shift operation where $T_n\varphi(x) = \varphi(x - n)$. Then the signal is reconstructed via the samples

$$\tilde{f}(x) = \sum_k f[k]T_{ka}\phi(x) \quad (2)$$

$\tilde{f}(x)$ is the approximation of $f(x)$ in the space V_ϕ , which is defined by

$$V_\phi = \{h(x) : h(x) = \sum c[n]T_{na}\phi(x), \sum |c[n]|^2 < \infty\} \quad (3)$$

When $\tilde{f}(x)$ is the projection of $f(x)$, $\phi(x)$ and $\varphi(x)$ are to satisfy the duality condition

$$\langle T_{ma}\varphi(x), T_{na}\phi(x) \rangle = \delta_{m,n} \quad (4)$$

Several criterias have been used to develop the sampling system; to name a few: Minimum Mean Square Error (MMSE), consistent as to the acquisition filters and minimax MSE [7], [8]. The common concern is for a given sampling kernel, how to design a sampling system such that the reconstructed signal satisfy certain optimality concern. However, there is no obvious solution on how to design a sampling system such that the signal can be perfectly reconstructed with minimum number of samples.

To achieve this goal, a proper sampling kernel as well as sampling rate should be tailored for the given signal. From (2), the reconstructed signal belongs to the space spanned by the shifts of $\phi(x)$. Therefore $\phi(x)$ should be chosen such that $f(x) \in V(\phi)$, only then $f(x)$ can be perfectly reconstructed. Furthermore, among all $\phi(x)$ that satisfy $f(x) \in V(\phi)$, different choice of ϕ results in different number of samples needed for perfect reconstruction [9]. Take the Multi-Resolution Analysis (MRA) using discrete wavelet for example. Assuming the mother wavelet is $\varphi(x)$, the wavelet frame is defined by

$$\varphi_j^{a,b} = (D_{a^j} T_{kb} \varphi)(x) = \frac{1}{a^{j/2}} \varphi\left(\frac{x}{a^j} - kb\right), \quad j, k \in Z \quad (5)$$

The MRA is carried out by the wavelet transform as

$$W_\varphi(f)(a, b) = \langle f, \varphi_j^{a,b} \rangle$$

To model the MRA system in a signal processing way: it is equivalent to pass the signal through a filter $\varphi(-x/a^j)$, then sample the output at rate $R = a^j b$ inasmuch as

$$\begin{aligned} f(x) * \varphi(-x/a^j) \cdot \sum_k \delta(x - ka^j b) \\ = \int_{-\infty}^{\infty} f(m) \varphi(m/a^j - kb) dm = a^{j/2} \langle f, \varphi_j^{a,b} \rangle \quad (6) \end{aligned}$$

The a, b are chosen to satisfy a loose admissibility condition [10]. Let $V_j = \text{span}\{\varphi_j^{a,b}\}$, from the properties of wavelet we have $V_j \subset V_{j+1}$. If $f(x)$ can be completely characterized at level j , it can be done as well at level k , $\forall k \geq j$. For fixed a, b , there exists a minimum j such that $f(x)$ can be fully characterized, which gives minimum R . For different choices of a, b and φ , it is daunting to identify which combination would result in the ultimate minimum R .

The RI of the signal can be used as a indicator to choose an appropriate sampling kernel for the given signal. For a signal $f(x)$, on one hand, ρ_f indicates its rate of degree of freedom, which implies the number of unknowns per unit time we need to specify for $f(x)$. On the other hand, if $f(x)$ is available, we would be empowered to settle unknowns at the rate of ρ_f . Refer to (1), when a sample $f[k]$ is taken, the correspondent sampling kernel is uniquely identified at $T_{ka}\varphi(x)$, therefore each sample should enable us to solve a number of unknowns at rate ρ_φ . Consequently, the relationship of ρ_f and ρ_φ would suggest us on how to choose a suitable sampling kernel for

the given signal. Before we proceed to choose the sampling kernel such that the sampling system is of minimum sampling rate, we first look at the problem of calculating the RI of a general signal.

III. CALCULATION OF RATE OF INNOVATION

There are some intuitive requirements that ρ should satisfy.

- 1) $\rho \geq 0$
- 2) $\rho_f = \rho_{Af(x-k)}$, A is a constant.
- 3) $\rho_f = B\rho_{f(x/B)}$, B is a constant.

It is because as an intrinsic property of the signal, any scaling in magnitude or simple shift in the variable should not affect ρ . However, the scaling of variable will render linear effect on ρ . We propose a simple method to calculate the RI of a signal in the shift invariant space:

Proposition 1: For a continuous signal $f(x)$ in Hilbert space, the RI of $f(x)$ can be calculated by

$$\rho_f = 1/|T| \quad (7)$$

where T satisfies $R_f(kT) = 0, \forall k \in Z, k \neq 0$, R_f is the autocorrelation function of $f(x)$ defined by $R_f(\tau) = \int_{-\infty}^{\infty} f(x)f(x-\tau)dx$.

Obviously Proposition 1 calculates RI that satisfies the requirements listed in the beginning of this section. There are some other properties that are noteworthy:

- 4) *Orthogonality*

If in a sampling system, ϕ is used as the sampling kernel and $a = 1/\rho_\phi$ is used as the sampling rate, then the kernel and its shifts $\{T_{na}\phi\}_{n \in Z}$ form a basis for its span V_ϕ . It is straightforward from Proposition 1. Since

$$\langle T_{na}\phi, T_{ma}\phi \rangle = R_\phi[(m-n)a] = \delta_{m,n}$$

and $\{T_{na}\phi\}_{n \in Z}$ consist of a frame for the span V_ϕ , $\{T_{na}\phi\}_{n \in Z}$ is the basis for V_ϕ . As a result, among all the families have the same span, the number of samples needed to specify a signal in V_ϕ is minimum with $\{T_{na}\phi\}_{n \in Z}$ as refer to (2).

- 5) *Independency*

The samples obtained from (1) using $\{T_{na}\phi\}_{n \in Z}$ as above are independent from each other. By 'independent' we mean that if one sample $f[k_1]$ is changed to $\Delta f[k_1]$, the rest should not be affected.

From (2), assume the reconstructed signal is $\tilde{\Delta f}(x)$ after $f[k_1]$ being replaced by $\Delta f[k_1]$

$$\tilde{\Delta f}(x) = \sum_{k \neq k_1} f[k] T_{ka}\phi(x) + \Delta f[k_1] T_{k_1 a}\phi(x) \quad (8)$$

To reinsert $\tilde{\Delta f}(x)$ to the sampling system, substitute (8) into (1) and exchange order of sum and inner product

$$\begin{aligned} f[m] = \sum_{k \neq k_1} f[k] \langle T_{ka}\phi(x), T_{ma}\varphi(x) \rangle \\ + \Delta f[k_1] \langle T_{k_1 a}\phi(x), T_{ma}\varphi(x) \rangle \quad (9) \end{aligned}$$

It is observed that the duality constraint in (4) is a necessary and sufficient condition for the samples to be

independent. Since $\{T_{na}\phi\}_{n \in \mathbb{Z}}$ form a basis for V_ϕ , the dual operator of $\phi(x)$ is itself and (4) is satisfied.

We show through experimental result in Section V that Proposition 1 produces consistent results for some specific functions. We also illustrate how it can be applied to a general signal in shift invariant space.

IV. IMPLEMENT SAMPLING SYSTEM WITH MINIMUM SAMPLING RATE

As we discussed in Section II, the RI of a signal should suggest us on the choice of the sampling kernel. From (1) and (2), the sampling process should define a reversible process if $f(x) = \tilde{f}(x)$, or the signal is perfectly reconstructed. It implies that the amount of information contained in $f(x)$ is preserved in $f[k]$ and then recovered to $\tilde{f}(x)$. The entity of 'information' is considered in the framework of sampling process as the degree of freedom of a signal to be specified. To break up the sampling and reconstruction process, and model the sampling process (1) as a Markoff process. As a joint production of $f(x)$ and $\varphi(x)$, the amount of information contained in $f[k]$ is restricted by $I_{\min}(f(x), \varphi(x))$. I_{\min} denotes the amount of information of $f(x)$ or $\varphi(x)$, which is lower. In other words, the degree of freedom of $f[k]$, or the RI of $f[k]$, is restricted by the lower of ρ_f and ρ_φ . Therefore, the sampling kernel chosen for a given signal to have minimum sampling rate should follow

Proposition 2: To sample a signal $f(x)$ of Rate of Innovation ρ_f , the sampling kernel $\varphi(x)$ of RI ρ_φ such that $\rho_\varphi \geq \rho_f$ should be used for perfect reconstruction. The minimum sampling rate is achieved when $\rho_\varphi = \rho_f$ and the sampling rate is $a = 1/\rho_\varphi$.

proof: Firstly, from Property (2), $\rho_\varphi = \rho_{T_{na}\varphi}$. The RI of the signal remains constant ρ_f throughout the entire sampling process. To reconstruct the signal is to reconstruct every piece of signal in the interval $[ka, (k+1)a]$. For every sample $f[k]$, the position of the sample determines the correspondent $T_{ka}\varphi$ and therefore is able to settle $N_1 = \rho_\varphi \cdot a$ unknowns. Similarly, during a time interval of a the number of unknowns of $f(x)$ is $N_2 = \rho_f \cdot a$. To eliminate all unknowns for $f(x)$ in $[ka, (k+1)a]$, we need $N_1 \geq N_2$, or $\rho_\varphi \geq \rho_f$.

Secondly, from Property (4), if the kernel φ of RI ρ_φ is used, when $a = 1/\rho_\varphi$, the family $\{T_{ka}\varphi(x)\}_{k \in \mathbb{Z}}$ consists of the basis for its span V_φ . For all signals $f(x) \in V_\varphi$, sampling using $\{T_{ka}\varphi(x)\}_{k \in \mathbb{Z}}$ results in the minimum number of samples.

We show in Section V how the classic Shannon's sampling theory and other existing sampling theories can be fitted into our design of sampling system of minimum sampling rate.

V. EXAMPLES

A. Calculation of RI

1) *sin(Bx):* Let $f(x) = \text{sinc}(Bx)$ and $g(x) = f(-x)$. Therefore $G(\omega) = F(-\omega)$. Let $z(x) = f(x) * g(x)$ we have

$$z(\tau) = \int_x \text{sinc}(Bx) \text{sinc}B(x - \tau) dt = R(\tau)$$

Let $Z(\omega)$ be the fourier response of $z(\tau)$,

$$Z(\omega) = F(\omega)G(\omega) = \begin{cases} \frac{1}{B^2} & |\omega| \leq \frac{B}{2} \\ 0 & \text{otherwise} \end{cases}$$

By IFT,

$$R(\tau) = z(\tau) = \frac{1}{B} \text{sinc}(B\tau)$$

Therefore, $R(\tau) = 0$ for $\tau = k/B$, $k \in \mathbb{Z}$, $k \neq 0$. The rate of innovation is given by $1/(1/B) = B$. It is consistent with the result that $\text{sinc}Bx$ function has B degree of freedom per unit time [1].

2) *Gaussian:* A Gaussian function $g(x)$ can be completely specified by two parameters, the variance σ^2 and the mean m . Therefore the totally degree of freedom is 2. Ideally, two samples should be able to describe $g(x)$. However, from Figure 2(a), it is observed that for any two samples at $t = a$ and $t = b$, two possible Gaussian functions can be reconstructed with means at $m1$ and $m2$ respectively. Therefore, in order to pinpoint a gaussian function, the extra information on the position of the two samples against the mean should be made available.

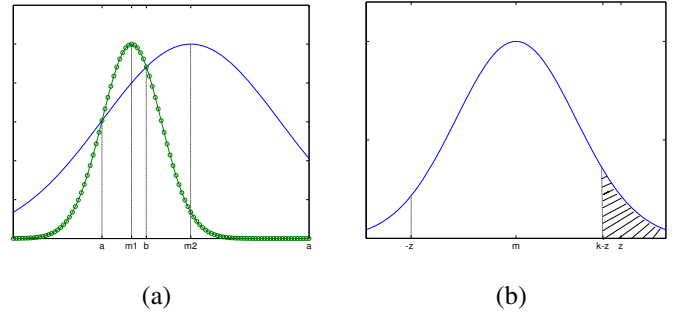


Fig. 2. (a) The reconstruction of Gaussian with two samples. Two Gaussian functions are possible if the position of the samples are not specified. (b) Relationship of k and z for a fixed threshold, meanings of symbols can be found in Section V-A.2.

To calculate RI for the gaussian function $g(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{\sigma^2}}$ using Proposition 1. After integration we obtain

$$R_g(k) = \int g(x)g(x-k)dx = \frac{1}{2\sigma\sqrt{\pi}} \exp(-k^2) \quad (10)$$

Since R_g is always positive, we choose a threshold P , $P \leq 1/\sqrt{2\pi}$ such that $R_g(k) \leq P$. The RI of gaussian is therefore $\rho_g = 1/k$ for a given threshold P .

Here we show that the ρ_g calculated using Proposition 1 is a truthful measure of degree of freedom for the gaussian functions. Using the same P to find the confidence range $[-z, z]$ such that $\int_{-z}^z g(x)dx \geq 1 - P$. Take the normal gaussian with $\sigma^2 = 1$ and $m = 0$ for example, some values of P and their corresponding k and z are listed in Table I. The entries of Table I is interpreted in this way: First, by comparing k and z for every same value of P , we find that $k \geq z$ and $k - z \geq 0$. Considering take the first sample at $-z$ and the next sample at $-z + k$, the area of the shaded

TABLE I

VALUES OF k AND z FOR DIFFERENT VALUES OF P IN CASE OF GAUSSIAN FUNCTION, $\sigma = 1$, $Q(x) = \int_0^x e^{-x^2/2} dx$

P	k	z	$ z - k $	$0.5 - Q(z - k)$
$1/\sqrt{(2\pi)}$	0	0	0	0.5
0.25	0.6951	0.6745	0.0206	0.4918
0.20	1.1729	0.8416	0.3313	0.3702
0.15	1.589	1.0364	0.5526	0.2903
0.10	2.0367	1.2816	0.7551	0.2251
0.05	2.6307	1.6449	0.9858	0.1621
0.01	3.6549	2.3263	1.3286	0.092
0.005	4.016	2.5758	1.4402	0.0749
0.001	4.7507	3.0902	1.6605	0.0484

area as in Figure 2(b) is denoted by $0.5 - Q(k - z)$ where Q is the error function defined by $Q(x) = \int_0^x e^{-x^2/2} dx$. For all P , $P \leq 0.5 - Q(k - z)$ which implies $k - z \leq z$. To summarize, for the same threshold P , it is always true that $z < k < 2z$. If we sample $g(x)$ at t and $t+k$ in the confidence range $[-z, z]$, the two samples obtained are always on different sides of the mean. As a result of that $g(x)$ can be reconstructed without ambiguity. Therefore the degree of gaussian function is $\rho_g = 1/k$, which is consistent with the result calculated using Proposition 1.

3) *Periodic Signal in Hilbert Square-integrable Space:* Now we consider a more general case where $f(t)$ is a complex-valued function of real argument t . Assume that $f(t)$ is piecewise continuous, periodic with period P and square integrable over an interval of P . The signal can be represented by its fourier series expansion where $f(t) = \sum_{n=-\infty}^{\infty} c_n e^{i\omega_n t}$, where $\omega_n = 2\pi n/P$. We would apply Proposition 1 to find out the RI of $f(t)$.

According to Prop 1, note that the autocorrelation defined for complex signal is $R_f(\tau) = \int f(t)f^*(t - \tau)dt$ where $*$ means the complex conjugate:

$$\begin{aligned} R_f(\tau) &= \int_{-P/2}^{P/2} \left[\sum_{n=-\infty}^{\infty} c_n e^{i\omega_n t} \right] \left[\sum_{m=-\infty}^{\infty} c_m e^{-i\omega_m (t-\tau)} \right] dt \\ &= \sum_{m,n=-\infty}^{\infty} c_n c_m e^{i\omega_m \tau} \int_{-P/2}^{P/2} e^{i2\pi(n-m)t/P} dt \end{aligned}$$

Since $\int_{-P/2}^{P/2} e^{i2\pi(n-m)t/P} dt = \delta_{n,m}$, (11) is reduced to $R_f(\tau) = \sum_{n=-\infty}^{\infty} c_n^2 e^{i\omega_n \tau}$. The RI is obtained by setting $R_f(\tau) = 0$. One possible solution is

$$\text{for all } n \quad e^{i\omega_n \tau} = \cos(\omega_n \tau) + i \sin(\omega_n \tau) = 0$$

$$\tau - \omega_0 = k \cdot \frac{P}{2n} \quad \omega_0 = \frac{1}{\omega_n} \tan^{-1}(-j) \quad k \in Z$$

Using Proposition 1, $T = \frac{P}{2n}$ and $\rho = 1/T$. As we notice from the result the RI depends on the frequency of the components of which $f(t)$ is made. The factor of 2 comes from 2 unknowns for each component, weight and frequency. It is consistent with the intuitive knowledge.

B. Design of sampling system with minimum sampling rate

1) *Shannon's sampling theory for bandlimited signals:* It is obvious that Shannon's sampling theory for bandlimited signals fits into our design of sampling system. From Section V-A.1, the function $\varphi(x) = \text{sinc}(Bx)$ has RI $\rho_\varphi = B$. It is also known that any signal $f(x)$, $F(\omega) = 0$ for $\omega > \pi B$ has a RI of $\rho_f = B$. Therefore, the optimal sampling kernel for $f(x)$ is $\varphi(x)$ such that $\rho_f = \rho_\varphi$ and the sampling rate is $a = 1/B$.

2) *non-bandlimited signals of finite RI:* One example is the sampling theory developed in [1]. Though the pulses are non-bandlimited, it still can be perfectly reconstructed by a *sinc* function of same RI. Another example is shown in [11]. It is shown that the pulse train can be perfect recovered using B-splines at the sampling rate equals to its RI.

VI. CONCLUSION

In this paper we consider the problem of calculating the Rate of Innovation for signals in the shift invariant Hilbert space. The RI of a signal can be used to identify a suitable sampling kernel as well as the sampling rate such that the signal can be perfectly reconstructed using minimum number of samples. We implement such sampling system using the RI calculated for the signal, and show through experimental results that some existing sampling theories can be viewed as examples of our implementation.

REFERENCES

- [1] Martin Vetterli, Pina Marziliano, and Thierry Blu, "Sampling signals with finite rate of innovation," *IEEE Trans. Signal Processing*, vol. 50, no. 6, pp. 1417–1427, June 2002.
- [2] P.P.Vaidyanathan, *Multirate Systems and Filter Banks*, prentice-Hall, Englewood Cliffs, New Jersey, 1993.
- [3] Michael Unser and Josiane Zerubia, "A generalized sampling theory without band limiting constrains," *IEEE Trans. Circuits Syst. II*, vol. 45, no. 8, pp. 959–969, Aug. 1998.
- [4] J.J. Clark, M.R. Palmer, and D. Lawrence, "A transformation method for the reconstruction of functions from nonuniformly spaced samples," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, no. 4, pp. 1151–1165, Oct. 1985.
- [5] Michael Unser, "Sampling - 50 years after Shannon," *Proc. IEEE*, vol. 88, no. 4, pp. 569–587, Apr. 2000.
- [6] Michael Unser, "Splines - a perfect fit for signal and image processing," *IEEE Signal Processing Mag.*, vol. 16, pp. 22–36, Nov. 1999.
- [7] Michael Unser and Akram Aldroubi, "A general sampling theory for non-ideal acquisition devices," *IEEE Trans. Signal Processing*, vol. 42, no. 11, pp. 2915–2925, Nov. 1994.
- [8] Yonina C Eldar and Michael Unser, "Nonideal sampling and interpolation from noisy observation in shift-invariant spaces," *IEEE Trans. Signal Processing*, vol. 54, no. 7, pp. 2636–2651, July 2006.
- [9] Beilei Huang and Edmund M-K Lai, "Non-bandlimited resampling of images," in *Proc. IEEE Int. Conf. Multimedia and Expo. IEEE*, July 2006, pp. 149–152.
- [10] Charles K. Chui, *An Introduction to Wavelets*, Wavelet analysis and its application. Academic press, 1992.
- [11] Beilei Huang, Edmund M-K Lai, and A.P Vinod, "Demodulation of UWB impulse radio signals using B-spline," in *Proc. IEEE Int. Conf. Communication systems*, 2006.