DISCRETE RANDOM SAMPLING THEORY

*Chenchi (Eric) Luo** *James H. McClellan*

Center for Signal and Information Processing
Georgia Institute of Technology

ABSTRACT

This paper proposes a new perspective on the relationship between the sampling and aliasing. Unlike the uniform sampling case, where the aliases are simply periodic replicas of the original spectrum, random sampling theory shows that the randomization of sampling intervals shapes the aliases into a noise floor in the sampled spectrum. New insights into both the Fourier random sampling problem and Compressive Sensing theory can be obtained using the theoretical framework of random sampling. This paper extends the theory of continuous time random sampling to deal with random discrete intervals generated from a clock. A key result is established to relate the discrete probability distribution of the sampling intervals to the power spectrum of the aliasing noise. Based on the proposed theory, a generic discrete random sampling hardware architecture is also proposed for sampling and reconstructing a class of spectrally sparse signals at an average rate significantly below the Nyquist rate of the signal.

***Index Terms*—** compressive sensing, random sampling

1. INTRODUCTION

In software defined radio (SDR), an ADC directly samples the signal from the antenna, which allows the front-end circuitry previously implemented in dedicated analog hardware to be moved into the digital domain. The wide spectrum range of the RF signal places a demanding performance requirement on the ADC. However, FCC reports on spectrum utilization reveal that even in the most densely packed urban areas the overall spectrum utilization rarely exceeds 35% at any one time. With the advent of Compressive Sensing (CS) [1], different ADC architectures[2, 3] have been proposed to reduce the sampling complexity by taking advantage of the sparse spectrum occupancy characteristic. Many of those architectures require specialized analog mixing circuits.

This paper develops a discrete random sampling theory that leads to a feasible architecture for discrete random sampling which can be integrated into existing standard ADC architectures without introducing extra mixing circuits. Using the proposed architecture, it is possible to sample and reconstruct a class of spectrally sparse signals at an average sampling rate below the Nyquist rate of the signal.

The investigation of this paper originated from the Fourier random sampling problem[4] that was studied before the emergence of the CS theory. The CS theory established a more general framework that justifies the effectiveness of the Fourier random sampling technique. On the other hand, since the proposed sampling interval in the Fourier random sampling problem follows a geometric probability distribution, a theory[5] that generalizes the relationship between the probability distribution of the sampling intervals and the spectrum of the sampled signal is also applicable. This paper generalizes the continuous time random sampling theory to deal with discretely distributed intervals which makes the theory more practical from an implementation perspective.

2. THE FOURIER RANDOM SAMPLING PROBLEM AND THE COMPRESSIVE SENSING THEORY

In the Fourier random sampling problem, given a signal $\mathbf{x} \in \mathbb{C}^N$, we are trying to find an optimal Fourier representation $\mathbf{x}_{\text{opt}}$ of K complex exponential terms to approximate $\mathbf{x}$. This clearly can be done by performing the fast Fourier transform (FFT) of $\mathbf{x}$ and locating the K largest terms. Gilbert et al. [4] showed that we can find a Fourier representation $\mathbf{x}^*$ by only sampling a subset $\mathcal{T} \subseteq [0, N-1]$ of $\mathbf{x}$ such that

$$||\mathbf{x} - \mathbf{x}^*||_2^2 \leq (1+\epsilon)||\mathbf{x} - \mathbf{x}_{\text{opt}}||_2^2, \quad (1)$$

where ϵ is an error bound parameter. The subset $\mathcal{T}$ can be set up by conducting independent Bernoulli trials on the index set $[0, N-1]$ with a fixed probability.

In Compressive Sensing[1], we can express $\mathbf{x}$ as

$$x[n] = \frac{1}{\sqrt{N}} \sum_{k=0}^{K-1} \alpha_k e^{j2\pi\omega_k n/N}, \quad \omega_k \subseteq [0, N-1]. \quad (2)$$

which can be written in matrix form, $\mathbf{x} = \mathbf{F}\boldsymbol{\alpha}$, where the elements of the discrete Fourier transform(DFT) matrix $\mathbf{F}$ are given by $F_{\omega,t} = \frac{1}{\sqrt{N}} e^{j2\pi\omega t/N}$, $\omega, t = 0, ..., N-1$, and $\boldsymbol{\alpha}$ only has K non-zero values at the normalized frequencies ω_k. The objective is to recover $\boldsymbol{\alpha}$ from the random samples of $\mathbf{x}$. The generation of the sampling subset $\mathcal{T}$ is exactly the same

*Supported in part by the Texas Instruments Leadership University grant.

as what is proposed in the Fourier random sampling problem. If the expected cardinality of $\mathcal{T}$ is M, then each sample on the uniform time grid is selected with probability M/N or discarded with probability $1-M/N$. Rudelson and Vershynin [6] showed that we can recover the sparse vector $\boldsymbol{\alpha}$ with high probability if $M = \mathcal{O}(K \log^4 N)$.

The sampling intervals in the above mentioned sampling scheme follow a geometric probability distribution. This paper investigates the relationship between the discrete probability distributions of the sampling intervals and the aliases, and justifies the feasibility of applying discrete random sampling to recover a sparse signal spectrum.

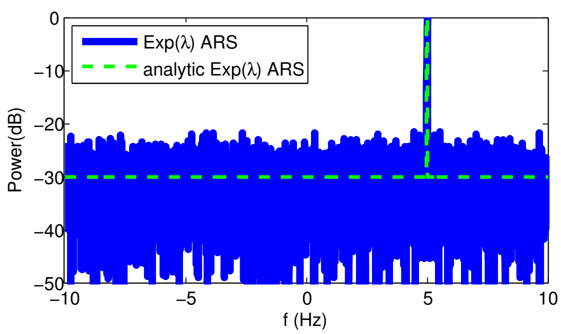

Fig. 1. Power spectra of a sampled sinusoidal signal with a frequency at 5 Hz for exponentially distributed ARS.

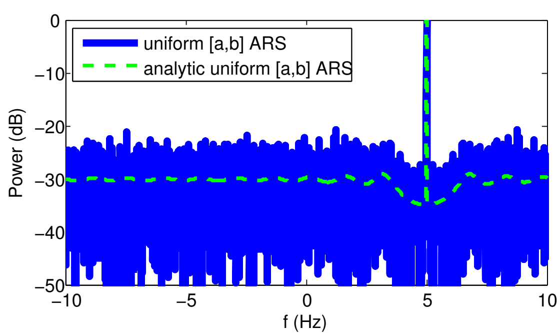

Fig. 2. Power spectra of a sampled sinusoidal signal with a frequency at 5 Hz for uniformly distributed ARS.

3. CONTINUOUS TIME RANDOM SAMPLING

Digital aliasing-free signal processing (DASP) was first mentioned by Shapiro and Silverman [7] in 1960. The key idea is to randomize the placement of sampling points so as to suppress aliasing. Beutler and Leneman [8, 9, 10, 5] published a series of papers on the theory of stationary point process and random sampling of random process in the late 1960s.

A random impulse process $s(t)$ is defined as

$$s(t) = \sum_{n=-\infty}^{\infty} \delta(t - t_n). \tag{3}$$

A random process $x(t)$ sampled by $s(t)$ can be written as $y(t) = x(t)s(t)$. If t_n is independent from $x(t)$, then

$$\Phi_y(f) = \Phi_x(f) * \Phi_s(f), \tag{4}$$

where $\Phi_y(f)$, $\Phi_x(f)$, $\Phi_s(f)$ are the power spectral densities (PSD) of $y(t)$, $x(t)$ and $s(t)$, respectively. When $t_n = nT$,

$$\Phi_s(f) = \frac{1}{T^2} \sum_{n=-\infty}^{\infty} \delta(f - \frac{n}{T}). \tag{5}$$

Therefore, aliases are periodic replicas of the signal spectrum under uniform sampling. The following theorem proposed in [5] generalized the analytic expression of $\Phi_s(f)$.

Theorem: Denote β as the average sampling frequency, $1/E[\tau_k]$, where τ_k are independently and identically distributed (i.i.d) intervals between samples. If the characteristic function of τ_k is $\psi_{\tau_k}(f)$, then

$$\Phi_s(f) = \beta \Re \left\{ \frac{1 + \psi_{\tau_k}(2\pi f)}{1 - \psi_{\tau_k}(2\pi f)} \right\}. \tag{6}$$

Since $\psi_{\tau_k}(0) = 1$, $\Phi_s(f)$ will have an impulse at $f = 0$.

3.1. Additive Random Sampling

In additive random sampling (ARS), the sampling time t_k is

$$t_k = t_{k-1} + \tau_k, \tag{7}$$

If the i.i.d interval τ_k follows an exponential distribution $\tau_k \sim \text{Exp}(\lambda)$, then

$$\Phi_s(f) = \lambda^2 \delta(f) + \lambda. \tag{8}$$

When τ_k is uniformly distributed $\tau_k \sim \text{Uniform}[a, b]$,

$$\Phi_s(f) = \begin{cases} P\left(\rho \frac{\sin((b-a)\pi f)}{(b-a)\pi f}, (b+a)\pi f\right) & f \neq 0 \\ \left(\frac{2}{a+b}\right)^2 \delta(f) & f = 0 \end{cases}, \tag{9}$$

where $P(r, \theta)$ is a Poisson kernel defined as:

$$P(r, \theta) = \frac{1 - r^2}{1 - 2r\cos\theta + r^2}, \tag{10}$$

Comparing (8,9) with (5), the $\delta(f)$ term will preserve the original signal spectrum, $\Phi_x(f)$, but the aliases no longer appear as frequency shifted replicas of $\Phi_x(f)$. Figs. 1 and 2 show the sampled power spectrum of an analytic sinusoidal signal with a frequency at 5 Hz. The average sampling frequency is $\beta = 3$ Hz, which is below the Nyquist rate. The number of samples is $M = 1024$. In both cases, the aliases are shaped into a noise floor. The shape of the noise floor depends on the distribution of the sampling intervals. The power of the aliasing noise floor is proportional to the power in the signal spectrum. Therefore, sparsity is required in the original signal spectrum to avoid very high aliasing noise power that would overwhelm the original signal spectrum.

4. DISCRETIZATION OF THE SAMPLING INTERVALS

In practice, it is difficult to implement continuously distributed intervals. The intervals are usually quantized onto a fixed time grid determined by a high-speed clock. Suppose the time grid has a uniform spacing of Δ, denote $\{\tau_k^q\}$ as the quantized time intervals, which are determined according to the rounding up criterion:

$$\tau_k^q = n\Delta, \text{if } (n-1)\Delta < \tau_k \leq n\Delta \qquad n \in \Omega, \tag{11}$$

where Ω is the corresponding feasible integer set. For example, if $\{\tau_k^q\}$ follows an exponential distribution, then $\Omega = \{1, 2, ...\}$. If $\{\tau_k^q\}$ follows a uniform distribution in $[0, T]$, then $\Omega = \{1, 2, ..., \lfloor \frac{T}{\Delta} \rfloor\}$. The time quantization of τ_k results in a periodic expansion of its characteristic function $\psi_{\tau_k}(f)$. Accordingly, $\Phi_s(f)$ also becomes periodic with a periodicity of $\frac{1}{\Delta}$. Therefore, we can ensure that the sampled signal is aliasing free only if $x(t)$ is bandlimited in $[-\frac{1}{2\Delta}, \frac{1}{2\Delta}]$. In other words, the minimum spacing rather than the average spacing of the sampling intervals determines the highest frequency that can be sampled without aliasing. However, the average sampling frequency is not completely free from the aliasing effect. Aliases in this case are not replicas of the original signal, but behave like spread spectrum noise, called aliasing noise here. The power of the aliasing noise is determined by the average sampling frequency.

5. THE DISCRETE RANDOM SAMPLING THEORY

This section extends the continuous-time random sampling theory to the discrete sampling interval case, and establishes a key result (18) that relates the discrete probability distribution of the sampling intervals with the power distribution of the aliasing noise.

If we denote Prob.$\{\tau_k^q = n\Delta\}$ as $p[n]$, then

$$\sum_{n\in\Omega} p[n] = 1. \tag{12}$$

We can define the discrete characteristic function of τ_k^q as the discrete time Fourier transform(DTFT) of the probability mass function (PMF) $p[n]$

$$\psi_{\tau_k^q}(e^{j\omega}) = \sum_{n\in\Omega} p[n]e^{j\omega n}, \tag{13}$$

where the normalized frequency ω is related to the continuous-time frequency f via

$$\omega = 2\pi f \Delta. \tag{14}$$

Accordingly, define the normalized aliasing noise power of $s(t)$ as

$$\Phi_n(e^{j\omega}) = \Re\left\{\frac{1+\psi_{\tau_k^q}(e^{j\omega})}{1-\psi_{\tau_k^q}(e^{j\omega})}\right\}, \quad \omega \in (0, \pi]. \tag{15}$$

We leave out the case where $\omega = 0$, which corresponds to an impulse function in $\Phi_n(e^{j\omega})$. Integrate $\Phi_n(e^{j\omega})$ over $(0, \pi]$, we can show that

$$\int_{0^+}^{\pi} \Phi_n(e^{j\omega})d\omega = \left(1 - \frac{1}{\sum_{n\in\Omega} np[n]}\right)\pi. \tag{16}$$

Note that

$$E(\tau_k^q) = \Delta \sum_{n\in\Omega} np[n], \tag{17}$$

the above equation can be rewritten as

$$\boxed{\underbrace{\frac{1}{\pi}\int_{0^+}^{\pi} \Phi_n(e^{j\omega})d\omega}_{\text{avg. noise power}} + \underbrace{\frac{\Delta}{E(\tau_k^q)}}_{\text{normalized avg. } f_s} = 1.} \tag{18}$$

Equation (18) represents a fundamental tradeoff between the average sampling frequency and aliasing noise power. We can reduce the average aliasing noise power by increasing the normalized average sampling frequency, or vice versa. As an extreme case, when $E[\tau_k^q] = \Delta$, which is exactly the uniform sampling case with spacing Δ, the aliasing noise power will decrease to zero.

For the discrete random sampling scheme adopted in the Fourier random sampling problem, the sampling interval follows a geometric probability distribution

$$p[n] = p^{n-1}(1-p), \tag{19}$$

where p is the skip probability. It is easy to verify that

$$\Phi_n(e^{j\omega}) = p, \quad \omega \in (0, \pi]. \tag{20}$$

$$\text{and} \quad E(\tau_k^q) = \frac{\Delta}{1-p}. \tag{21}$$

Thus, we can decrease the average sampling rate by increasing the skip probability p, which will also raise the flat aliasing noise floor.

A generalization of the geometric probability distribution is the negative binomial distribution, which represents the probability of n successes in a sequence of independent Bernoulli trials until r failures occur. When $r = 1$, the negative binomial distribution degenerates into the geometric probability distribution. If the success probability is denoted as p, then we have

$$p[n] = \binom{n+r-2}{n-1} p^{n-1}(1-p)^r, \quad n = 1, ... \tag{22}$$

$$E(\tau_k^q) = \left(\frac{pr}{1-p} + 1\right)\Delta. \tag{23}$$

When the average sampling interval $E(\tau_k^q)$ is fixed, so is the average aliasing noise power. However, we have the freedom to shape the power distributions of $\Phi_n(e^{j\omega})$ over $(0, \pi]$ by tuning p and r. Fig. 3 shows an example of different aliasing noise PSDs relative to the original signal frequency with the same average sampling interval $E[\tau_k^q] = 5.2\Delta$. As r increases, there will be a deeper dip around $\omega = 0$, where the original signal frequency resides. Therefore, the aliasing noise has less impact under the signal frequency when r increases. The reduced aliasing noise power is shaped into other frequency bands. In other words, we can distribute the power distribution of the aliasing noise by designing an appropriate probability mass function for the random discrete intervals.

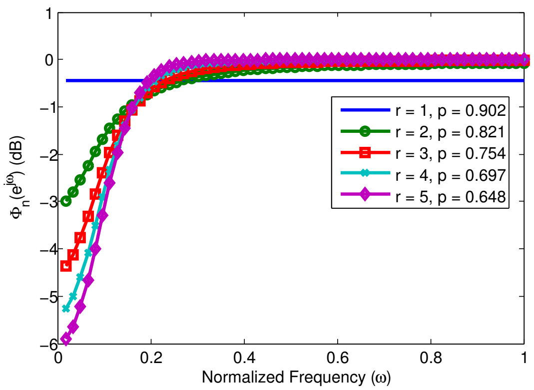

Fig. 3. The aliasing noise power spectrum $\Phi_n(e^{j\omega})$ (dB) relative to the original signal frequency ($\omega = 0$) for negative binomial distributed ARS, $E[\tau_k^q] = 5.2\Delta$.

6. A GENERIC DISCRETE ADDITIVE RANDOM SAMPLING ARCHITECTURE

Based on the discrete random sampling theory, Fig. 4 shows the block diagram of a generic ARS scheme. The random number generator produces a discrete interval $\tau_k^q = n_k\Delta$ according to a prescribed PMF. The input signal $x(t)$ is sampled at t_k and quantized to $x^q(t_k)$. Finally, $x^q(t_k)$ and the interval integer n_k are both encoded and packaged together as a pair.

Amplitude quantization (number of bits) is an important parameter for ADCs. If the amplitude of each sample $x(t_k)$ can be quantized to full precision within the minimal interval Δ, then the amplitude quantization noise can still be treated as uniformly distributed additive noise. However, a fundamental tradeoff in almost all ADC hardware is the tradeoff between amplitude resolution and sampling speed. Allowing each sample to reach its full amplitude precision increases the minimal time interval Δ and reduces the frequency coverage. A sampling model that leverages this tradeoff was proposed in [11].

7. SIGNAL RECONSTRUCTION FROM THE RANDOM SAMPLES

The clocked time quantization in the ARS architecture makes it possible to calculate the power spectrum using an FFT by

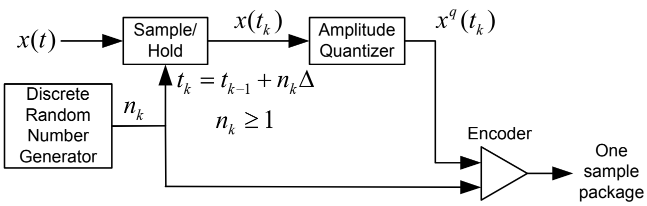

Fig. 4. The block diagram of a generic ARS scheme

replacing missing values with zeros. After we have collected M quantized samples with $\{x^q(t_k), n_k\}, k = 1, ..., M$ and $N = \sum_{k=1}^{M} n_k$, we can insert zeros in between each sample according to n_k as shown in Fig. 5. If we denote the zero-inserted signal vector as $\bar{\mathbf{x}}$, then we can calculate the normalized power spectrum an N-point FFT

$$\mathbf{p} = \frac{1}{M^2}\left|\text{FFT}_N\{\bar{\mathbf{x}}\}\right|^2, \tag{24}$$

Since the minimal time interval is Δ, the frequency grid spacing is $\frac{1}{N\Delta}$ Hz. If we are only interested in the detection

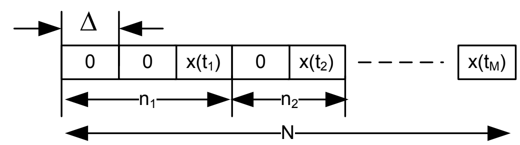

Fig. 5. Interval zero insertion for ARS samples.

of certain frequency components from the random samples, calculating the power spectrum is sufficient. In some applications, it is desirable to reconstruct the randomly sampled signal onto a fine uniform time grid so that it can be further processed by classic DSP systems. However, not all bandlimited signals can be recovered from the random samples because the aliasing noise introduced by the random sampling process could overwhelm the original signal spectrum. Since the aliasing noise power is proportional to the spectrum occupancy of the original signal, only those signals with a sparse spectrum occupancy can be successfully reconstructed. Three factors: aliasing noise, spectral leakage, and amplitude quantization noise make the sampled signal not perfectly sparse in the frequency domain. Generic reconstruction algorithms in Compressive Sensing such as OMP[12] or CoSaMP[13] was customized in [14] to deal with this specific sparse spectrum reconstruction problem.

8. REFERENCES

[1] E.J. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *Information Theory,*

IEEE Transactions on, vol. 52, no. 2, pp. 489 – 509, Feb. 2006.

[2] J.A. Tropp, J.N. Laska, M.F. Duarte, J.K. Romberg, and R.G. Baraniuk, "Beyond nyquist: Efficient sampling of sparse bandlimited signals," *Information Theory, IEEE Transactions on*, vol. 56, no. 1, pp. 520 –544, Jan. 2010.

[3] M. Mishali and Y.C. Eldar, "From theory to practice: Sub-nyquist sampling of sparse wideband analog signals," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 4, no. 2, pp. 375 –391, Apr. 2010.

[4] A. C. Gilbert, S. Guha, P. Indyk, S. Muthukrishnan, and M. Strauss, "Near-optimal sparse fourier representations via sampling," in *Proceedings of the thiry-fourth annual ACM symposium on Theory of computing*, New York, NY, USA, 2002, STOC '02, pp. 152–161, ACM.

[5] F. J. Beutler and O. A. Z. Leneman, "The spectral analysis of impulse processes," *Information and Control*, vol. 12, pp. 236–258, 1968.

[6] M. Rudelson and R. Vershynin, "Sparse reconstruction by convex relaxation: Fourier and gaussian measurements," *Information Sciences and Systems, 2006 40th Annual Conference on*, pp. 207 –212, Mar. 2006.

[7] H. S. Shapiro and R. A. Silverman, "Aliasing-free sampling of random noises," *SIAM J. Appl. Math.*, vol. 8, pp. 225–236, 1960.

[8] F. J. Beutler and O. A. Z. Leneman, "The theory of stationary point processes," *Acta Math.*, vol. 116, pp. 159–197, 1966.

[9] F. J. Beutler and O. A. Z. Leneman, "Random sampling of random process: Stationary point processes," *Information and Control*, vol. 9, pp. 325–344, 1966.

[10] O. A. Z. Leneman, "Random sampling of random process: Impulse processes," *Information and Control*, vol. 9, pp. 347–363, 1966.

[11] Chenchi Luo and James H. McClellan, "Compressive sampling with a successive approximation adc architecture," in *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on*, May 2011, pp. 3920 –3923.

[12] J.A. Tropp and A.C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *Information Theory, IEEE Transactions on*, vol. 53, no. 12, pp. 4655 –4666, Dec. 2007.

[13] D. Needell and J. A. Tropp, "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples," *Applied and Computational Harmonic Analysis*, vol. 26, no. 3, pp. 301–321, Apr. 2008.

[14] Chenchi Luo, *Non-uniform Sampling: Algorithms and Architectures*, Ph.D. thesis, Georgia Institute of Technology, 2012.