Kolmogorov Complexity of Signals with Finite Rate of Innovation

Subhas K. Ghosh, Viswanath Ganapathy, Chandrashekhara Thejaswi and Ranjeet K. Patro
Honeywell Technology Solutions Laboratory
151/1, Doraisanipalya, Bannerghatta Road,
Bangalore, India, 560076
e-mail: subhas.kumar@honeywell.com

***Abstract*— We consider the class of signals that has finite rate of innovation. There are few fundamental questions associated with such class of signals. One such question is about a generic characterization. In this work we show that the signals which has finite rate of innovation are also algorithmically compressible. This characterization is useful because algorithmic characterization allow us to view such classes of signal from information theoretic point of view. In this work we relate the parametric representation of such signals to their string equivalent, which can be generated by a short program on universal computer or has low Kolmogorov complexity. We show how this characterization applies to signals that have finite number of freedom per unit time, with case studies.**

I. Introduction

A physical process as a continuous phenomena can be represented as a continuous time signal $x(t)$. Corresponding to this signal $x(t)$ we can define $y(t) = x(t) * \tilde{h}(t)$ as the filtered version of $x(t)$ with $\tilde{h}(t) = h(-t)$ as the convolution kernel. To observe such continuous phenomena a uniform sampling with sampling interval T produces $y(nT)$, given by: $y(nT) = \langle h(t-nT), x(t)\rangle = \int_{-\infty}^{\infty} h(t-nT)x(t)dt$ and the intermediate signal can be presented as: $y_s(t) = \sum_{n\in\mathbb{Z}} y(nT)\delta(t-nT)$ It is essential that we have a faithful representation of $x(t)$ in the sample $y(nT) = \langle h(t-nT), x(t)\rangle$, so that we can reconstruct $x(t)$ and understand the physical phenomena. An alternative question will be, if we do not have a faithful representation of $x(t)$, then to what extent we can reconstruct $x(t)$? This question is fundamental, and it is known that if $x(t)$ has bandwidth $B = 2\omega_m/2\pi$, $B = 1/T$ samples are sufficient to reconstruct $x(t)$ as: $x(t) = \sum_{-\infty}^{\infty} x(nT)\sin(\pi(\frac{t}{T}-n))/\pi(\frac{t}{T}-n)$ Any real band limited signal can be seen as having $1/T$ degrees of freedom per unit time, which is the number of samples per unit time that sufficiently specify it, which is also known as the rate of innovation of the signal $\rho = \frac{1}{T} = \frac{\omega n}{\pi}$.

The signals with finite rate of innovation are of special interest. Given a band limited signal with finite rate of innovation we can sample it with ρ samples per unit time with uniform sampling rate and reconstruct easily. This is not so when we consider a larger class of signals though. In [VMB02] it was shown that stream of Diracs, nonuniform splines, and piecewise polynomials, can be sampled uniformly at (or above) the rate of innovation using appropriate kernel and then can be perfectly reconstructed. In [RMK02] authors have validated sampling based on finite rate of innovations for signals in noise. For bandwidth expanding communication systems it is shown that we can design low complexity receivers based on finite rate of innovations. A valid application of the same was presented in [KRV02], [IMV03]. All these work raises a fundamental question:

Question 1 *Are there any generic characterization of signals with finite rate of innovation, which can be uniformly sampled and then be acceptably reconstructed?*

In this work we provide a positive answer to this question by relating signals with finite rate of innovation to Kolmogorov complexity. We show that a signal with finite rate of innovation can be uniformly sampled and then be acceptably reconstructed if and only if they are algorithmically compressible. Where, we relate the parametric representation of such signals to their string equivalent, which can be generated by a short program. This characterization is useful because algorithmic characterization allow us to view such classes of signal from information theoretic point of view [CL78], [GV].

II. PRELIMINARIES

Throughout this paper we fix our alphabet as Σ. By string we denote the elements of Σ^*. We denote the length of string x by $|x|$. Let $\mathcal{U}$ be an universal computing machine capable of computing partial recursive functions of the form $\varphi : \Sigma^* \rightarrow \Sigma^*$, and $\mathcal{U}(P)$ denote the output of the machine when presented with program P, if it is defined otherwise machine fails and halts generating Λ. We can think of P as another string in Σ^* and $|P|$ denotes the length of such program. We consider that the program P is self delimiting.

DEFINITION 1 ***The Kolmogorov (Algorithmic) complexity***: *Kolmogorov complexity of a string $x \in \Sigma^*$, with respect to universal machine $\mathcal{U}$ is: $K_{\mathcal{U}}(x) = \min_{P:\mathcal{U}(P)=x} |P|$ The minimum length input program P that generates x and halts on $\mathcal{U}$. Above definition assumes that $\mathcal{U}$ is just provided with the program P, but it has no idea on the length of x. If the length of x is known then we can define the conditional Kolmogorov complexity as: $K_{\mathcal{U}}(x|\,|x|) = \min_{P:\mathcal{U}(P,|x|)=x} |P|$*

DEFINITION 2 ***Signals with finite rate of innovation***: *We say a signal with a* finite rate of innovation*[VMB02] is a signal which has a parametric representation $x(t) = \sum_{n\in\mathbb{Z}} \sum_{r=0}^{R} c_{nr}\phi_r\left(\frac{t-t_n}{T}\right)$ with arbitrary shifts, and assuming a set of functions $\{\phi_r(t)\}_{r=0,\ldots,R}$ are known, and there exists a counting function $C_x(t_a, t_b)$ defined over interval $[t_a, t_b]$, used to count degrees of freedom c_{nr} at time instant t_n, where we define the rate of innovation ρ as: $\rho = \lim_{\tau\rightarrow\infty} \frac{1}{\tau} C_x(-\frac{\tau}{2}, \frac{\tau}{2})$ with ρ being finite.*

DEFINITION 3 ***Sufficient characterizer***:*Let $h_B(t)$ be a sampling kernel, with $B \geq \rho$, where ρ is the rate of innovation, and sample $(h_B * x)(t)$ at N uniform locations $t = nT, n = 0, \ldots, N-1$, then samples $y_n = \langle h(t-nT), x(t)\rangle, n = 0, \ldots, N-1$ is a sufficient characterize if and only if for a universal reconstruction algorithm U_R, y_n provides a unique solution for $x(t)$.*

DEFINITION 4
Shift invariant spaces: We call a $\mathcal{S}$ a (finitely generated) shift invariant space, if $\mathcal{S} = \sum_{n\in\mathbb{Z}} \sum_{r=0}^{R} c_{nr}\phi_r\left(\frac{t}{T} - n\right)$ where,$\{\phi_r(t)\}_{r=0,\ldots,R}$ are the generating functions of $\mathcal{S}$ and R is called the length of $\mathcal{S}$, denoted as $len(\mathcal{S})$.

DEFINITION 5 ***Signals with finite rate of innovation***: *Let $\mathcal{X}$ denote class of signals in Hilbert space $\mathcal{H}$. We call $\mathcal{X}$ as class of signals with finite rate of innovation if $\mathcal{X} = \cup_{i\in\mathcal{I}}\mathcal{S}_i$, where $\{\mathcal{S}_i\}_{i\in\mathcal{I}}$ is a collection of shift invariant subspaces with $len(\mathcal{S}_i) = R_i$, and $i \in \mathcal{I}$ are indices of the subspace, $\mathcal{I} \subseteq \mathbb{R}^m$. This corresponds to the definition given in 2, where set of functions $\{\phi_r(t)\}_{r=0,\ldots,R_i}$ are the generating functions of the shift invariant subspaces $\mathcal{S}_i$.*
Intuitively, the signal is not *smooth*, as it spans only over finite number of subspaces per unit time.

III. CHARACTERIZATION USING KOLMOGOROV COMPLEXITY

To establish any meaningful relation with a sampled signal, notice that we have to relate the string generated by the universal machine with the sample of the form $y_n = \langle h(t-nT), x(t)\rangle, n = 0, \ldots, N-1$. Given the universal computer can output at most one symbol in unit time, we shall consider a time bounded Kolmogorov complexity captures the notion of a sampled ensemble. First we will assume that universal computer can output at most one symbol in unit time on its one-sided output tape. Thus the generated sequence of the universal machine $x \in \Sigma^*$ from $\mathcal{U}$ on input P can be thought of a sequence $x_1\#x_2\#\ldots\#x_n$, where each x_i has been generated based on some specification given in P, or the partial recursive function φ_P computed by $\mathcal{U}$ as specified by P, with one x_i at a time, and $\#$ is a delimiter and $\# \notin \Sigma$. We shall also consider that $\forall 1 \leq i, j \leq n, |x_i| = |x_j|$. We will briefly state one useful result and refer [HR67] for details:

THEOREM 1
Composition is effective, i.e. for all n and p there is a primitive recursive function $Comp(i_1, \ldots, i_n, j)$ such that if $j, i_1, \ldots, i_n$ are indices for the n-ary function g and the p-ary functions $h_1, \ldots, h_n$ respectively, then $Comp(i_1, \ldots, i_n, j)$ is an index for the p-ary composition $f = g(h_1, \ldots, h_n)$.
We will define a notion similar to finite rate of innovation for the string generated by the universal machine.

DEFINITION 6 ***Finite rate of variation***: *Let $\mathcal{U}$ be an universal machine computing $x = x_1\#x_2\#\ldots\#x_n$ on input P, and $\forall 1 \leq i, j \leq n, |x_i| = |x_j| = m$. Assume, there exist a counting function $C_x(a, b)$ defined over indexes of x, $[a, b], 1 \leq a, b \leq n$, used to count $d_{i,j}(x_i, x_j) = \|\{k|x_{i,k} \neq x_{j,k}\}\|$, where $x_{i,k}$ is the k^{th} alphabet of x_i, for each $a \leq i, j \leq b$.*

we say that x has finite rate of variation if and only if $\nu = \sum_{\forall a,b\in\{0,\dots,n\}\mid a<b} \frac{C_x(a,b)}{m\binom{n}{2}} < 1$ In fact it is not necessary that $\forall 1 \le i,j \le n, |x_i|$ be same as $|x_j|$, and we can relax this assumption, however it is useful to make this assumption for easier exposition.

Intuitively, the way a physical signal is composed by a set of functions $\{\phi_r(t)\}_{r=0,\dots,R}$, the program on universal machine is also composed of family of recursive functions. Now, we state some of the known relations between entropy of a sequence $H(X)$ and Kolmogorov complexity, and then compressibility of a sequence with the same [CT91].

THEOREM 2

***Entropy and Kolmogorov complexity**: Let X_n is a probability distribution on some discrete set Σ, according to the probability mass function $f(x), x \in \Sigma$. Let, $f(x^n) = \prod_{i=1}^n f(x_i)$, then there exists a constant ϵ such that: $H(X) \le \frac{1}{n}\sum_{x^n} f(x^n)K_{\mathcal{U}}(x^n|n) \le H(X) + \frac{|\Sigma|\log n}{n} + \frac{\epsilon}{n}$ for all n.*

THEOREM 3

***Compression and Kolmogorov complexity**: A string x is compressible if and only if $K_{\mathcal{U}}(x) < |x|$*

THEOREM 4

Let $\mathcal{U}$ be a universal machine, that on input $\langle P, n\rangle$ generates $\mathcal{U}(P,N) = x \in \Sigma^*$, as a sequence $x_1\#x_2\#\dots\#x_n$, then x has finite rate of variation if and only if x is compressible.

PROOF: Let $\mathcal{U}$ be an universal machine computing $x = x_1\#x_2\#\dots\#x_n$ on input P, and $\forall 1 \le i,j \le n, |x_i| = |x_j| = m$. We will first assume that x has finite rate of variation $\nu \le 1$. Consider the following program P:

procedure P **input** $\langle \nu, m, n\rangle$
begin
Generate, in lexicographic order, all sequences of form $x = x_1\#x_2\#\dots\#x_n$
Keep two conditional variables a and b with range $\{1,\dots,n\}$ and $a < b$
Generate $x = x_1\#x_2\#\dots\#x_n$ in n steps, by generating x_i in step i with $|x_i| = m$
In step i $\forall a,b \in \{1,\dots,i\}$ i.e. $\binom{i}{2}$ options of a,b
Ensure that x_i maintains ν rate of variation
end.

We can compute the $|P|$, the length of the program as follows. We need to maintain variable a and b in every step, take $\binom{i}{2}$ different values, and for each value check $(b-a)m$ places. Thus it is enough for the program to keep variables a, b, a variables to check all $\binom{i}{2}$ options has been generated in step i, a variable to check the size of x_i, and a variable for ν. $K_{\mathcal{U}}(x) \le |P| = c + \underbrace{4\log n}_{a,b} + \underbrace{2\log m}_{|x_i|} + \underbrace{\log\binom{n}{2}}_{options} + \underbrace{2\log\nu}_{rate} \le c + 8\log n + nH_0(\frac{2}{n})$ Considering $\binom{n}{k} \le 2^{nH_0(k/n)}$, which follows from Stirling's formula, and introducing $H_0(p) = -p\log p - (1-p)\log(1-p)$ as binary entropy. Thus x of length mn is surely compressible.

Now consider we have $K_{\mathcal{U}}(x) < |x|$. The program can be P^* :"generate a mn length string from x'", where we assume $|x'| = |x| - \delta$, where δ is some positive constant. Now it is easy to see that to *expand* x', while to keep x compressible, at least δ portion of x' will appear in x at least twice, thus x will have $\nu < 1$. ■

DEFINITION 7 *For class of signal $\mathcal{X}$ in the $\mathcal{H}$ signal space, we define an orthogonal projection operator $\mathbf{P}_{\mathcal{S}} : \mathcal{H} \longmapsto \mathcal{S}$ from $\mathcal{H}$ to one of its subspace $\mathcal{S}$, as a representation of $\mathcal{X}$, if any signal $\mathbf{x} \in \mathcal{X}$ can be uniquely determined by its image $\mathbf{P}_{\mathcal{S}}\mathbf{x}$, i.e. $\mathbf{P}_{\mathcal{S}}\mathbf{x}_1 = \mathbf{P}_{\mathcal{S}}\mathbf{x}_2$ if and only if $\mathbf{x}_1 = \mathbf{x}_2$.*

THEOREM 5

[LD04] For class of signal $\mathcal{X}$ in the $\mathcal{H}$ signal space having finite rate of innovation, i.e. $\mathcal{X} = \cup_{i\in\mathcal{I}}\mathcal{S}_i, \mathcal{S}_i \subseteq \mathcal{H}$, and these subspaces are invariant under shifts, and $i \in \mathcal{I}$ are indices of the subspace, $\mathcal{I} \subseteq \mathbb{R}^m$. A projection operator $\mathbf{P}$ satisfying definition above if and only if for all $\tilde{\mathcal{S}}_{ij} \triangleq \mathcal{S}_i + \mathcal{S}_j$ with $(i,j) \in \mathcal{I}\times\mathcal{I}$, $\mathbf{P}$ is an injective operator from $\tilde{\mathcal{S}}_{ij}$ onto $\mathbf{P}\tilde{\mathcal{S}}_{ij}$

For a signal with finite rate of innovation, $\dim(\mathcal{S}) \ge \max_{(i,j)\in\mathcal{I}\times\mathcal{I}} \dim(\tilde{\mathcal{S}}_{ij})$ is a necessary condition. At the same time this allows us to well represent the signal with $\dim(\mathcal{S})$ elements. An equivalent condition for the above is for all $\tilde{\mathcal{S}}_{ij}$ with $(i,j) \in \mathcal{I}\times\mathcal{I}$ the following $R \times N$ matrix $\mathbf{G} = \langle g_i, k_j\rangle$ has full column rank, where $\{g_i\}_{i=1}^N$ are the basis vectors of $\tilde{\mathcal{S}}_{ij}$ and $\{k_i\}_{i=1}^R$ are the basis vectors of $\mathcal{S}$.

THEOREM 6

Finite rate of innovation $\Leftrightarrow$ finite rate of variation.

PROOF: We provide a sketch of proof to show this. Consider a signal $\mathbf{x} \in \mathcal{X}$, where $\dim(\mathcal{S})$ is sufficient to represent signals in class $\mathcal{X}$ as $\cup_{i\in\mathcal{I}}\mathcal{S}_i$. Since we have $\dim(\mathcal{S})$ real numbers sufficient to represent $\mathbf{x} \in \mathcal{X}$, we define a index for such real numbers as a quantization function $Qnt : \mathbb{R} \to \mathbb{Z}$, range of Qnt is finite. We allow a string $x = x_1 \# x_2 \# \ldots \# x_n$ to be generated by $\mathcal{U}$ using $\{\varphi_i\}$ having range in $\subseteq range(Qnt)$. We need to show that x generated thus will have finite rate of variations. We can think each x_i as a vector of the form $x_i = \langle x_{1,i}, x_{2,i}, \ldots, x_{m,i} \rangle$. Since each of $x_{j,i} \in S \subseteq range(Qnt)$, we can think each x_i is generated from a m-dimensional subspace. Thus we can represent the class of strings of type x as $\mathcal{Y} = \cup_{i \in \mathcal{I}} S_i$, where each S_i is an m-dimensional subspace and $S_i \subseteq range(Qnt)$. We can think of partial recursive functions as generating functions for each such subspace. Here again finite rate of variation means that x spans only over finite number of subspaces. Thus we conclude that Finite rate of innovation $\Rightarrow$ finite rate of variation.

For the other side observe that we can represent the signals with finite rate of innovation by equivalent condition as for all $\tilde{\mathcal{S}}_{ij}$ with $(i,j) \in \mathcal{I} \times \mathcal{I}$ the $R \times N$ matrix $\mathbf{G}$ with full column rank, where $\{g_i\}_{i=1}^{N}$ are the basis vectors of $\tilde{\mathcal{S}}_{ij}$ and $\{k_i\}_{i=1}^{R}$ are the basis vectors of $\mathcal{S}$. We can take $\{g_i = f(t - t_i)\}_{i=1}^{R}$ with $0 \leq t_1 < t_2 < \ldots < t_R < T$, and $\{k_i = f'(t)\}_{i=1}^{R}$, where f and f' are some generating function ranging over $[0, T]$, then we have $\mathbf{G} = \left[\left\langle f(t - t_i), f'^j(t) \right\rangle\right]$, and each $\left\langle f(t - t_i), f'^j(t) \right\rangle$, represent the inner product. We claim that there exist a finite program P that can implement $\mathbf{G}$. Thus the output of the program will have strings generated with finite rate of variation and Finite rate of innovation $\Leftarrow$ finite rate of variation. This completes the proof. ■

COROLLARY 7

A signal with finite rate of innovation is Kolmogorov compressible.

IV. COMPUTATIONAL COMPLEXITY

The question on complexity is related to hardness of representing these signals in this case. To make this precise we define the notion of $\mathbf{AC}^0$ reduction. Following the presentation of [CSV84] a Boolean function f is $\mathbf{AC}^0$-reducible to the set of Boolean functions $\mathcal{F}$, i.e. $f \in \mathbf{AC}^0(\mathcal{F})$, if there is a constant depth, polynomial size circuit $\mathbf{C}$ with basic operations $\mathcal{F}$ and output f. We call A Boolean function $f \in \mathcal{B}_n$ is symmetric if

$$f(x_1, x_2, \ldots, f_n) = f(x_{\pi(1)}, x_{\pi(2)}, \ldots, f_{\pi(n)}) \tag{1}$$

for every permutation π in the group S_n of permutations on n letters. It is known that Every symmetric function is in $\mathbf{AC}^0(\{1, \oplus, \wedge, MAJ\})$, where MAJ is Majority: outputs 1, if and only if the majority of the inputs is 1.

Define the Boolean matrix product $C = (c_{i,j}) = A \cdot B$, for $n \times n$ matrices A, B by

$$c_{i,j} = \bigvee_{k=1}^{n} a_{i,k} \wedge b_{k,j}, 1 \leq i,j \leq n, a_{i,k}, b_{k,j} \in \{0,1\} \tag{2}$$

Thus a constant depth, $O(n^3)$ size, unbounded fan-in Boolean circuit computing matrix product.

Few more notations, $\mathbf{NC}^k$ denotes the family of all functions $f : \{0,1\}^* \to \{0,1\}$, such that f is computed by a family $\langle \mathbf{C}_n \rangle_{n \in \mathbb{N}}$ of fan-in 2 Boolean circuits, where depth of the circuit is $O(\log^k n$ and size of the circuit is $n^{O(1)}$. We define $\mathbf{AC}^k$ in the same way, with unbounded fan-in circuits. Thus $\mathbf{NC}^k \subseteq \mathbf{AC}^k$ and $\mathbf{AC}^k \subseteq \mathbf{NC}^{k+1}$.

PROPOSITION 8

Let $\mathcal{X}$ be class of signals with finite rate of innovation $\mathcal{X} = \cup_{i \in \mathcal{I}} \mathcal{S}_i$, where $\{\mathcal{S}_i\}_{i \in \mathcal{I}}$ is a collection of shift invariant subspaces with $len(\mathcal{S}_i) \leq R_i$, and $i \in \mathcal{I}$ are indices of the subspace, $\mathcal{I} \subseteq \mathbb{R}^m$. Let $\mathbf{x} \in \mathcal{X}$ be a signal, and let $L(Samp)$ represent the language capturing the problem of sampling signals with finite rate of innovation, then $L(Samp) \in \mathbf{NC}^1$

PROOF: We first state the following:

- Integer addition is in $\mathbf{AC}^0$.
- Integer multiplication is in $\mathbf{NC}^1$.
- Integer division is in $\mathbf{NC}^1$.

Now recall that we can represent the signals with finite rate of innovation by equivalent condition as for all $\tilde{\mathcal{S}}_{ij}$ with $(i,j) \in \mathcal{I} \times \mathcal{I}$ the following $R \times N$ matrix, which has full column rank, where $\{g_i\}_{i=1}^{N}$ are the basis vectors of $\tilde{\mathcal{S}}_{ij}$ and $\{k_i\}_{i=1}^{R}$ are the basis vectors of $\mathcal{S}$. We can take $\{g_i = f(t - t_i)\}_{i=1}^{R}$ with $0 \leq t_1 < t_2 < \ldots < t_R < T$, and $\{k_i = f'(t)\}_{i=1}^{R}$, where f and f' are some generating function ranging over $[0, T]$. Finally each entry in matrix $\mathbf{G}$ is $\left\langle f(t - t_i), f'^j(t) \right\rangle$, representing the inner product. We claim that to implement this inner product using addition, multiplication and division, we will be in $\mathbf{NC}^1$. Finally, to compute $\mathbf{Gx}$, we need at most $\mathbf{NC}^1$ circuit again. ■

V. Case study

Here we discuss one example illustrating the connection between Kolmogorov complexity and finite rate of innovations. Consider the class of signals that can be represented as $x(t) = \sum_{k=1}^{K} c_k \delta(t - t_k)$ within each period T. Here t_k can be chosen randomly from $[0, T)$. For the case where t_k is an arbitrary real number in the interval $[0, T)$ we assume that the rational approximation is used to represent t_k. The c_k's can also be considered as a real number.

Ultra wide band signaling using pulse position modulation can be represented as a weighted sum of Diracs. For these classes of signals the innovation rate is lower than the bandwidth of the signal and therefore there is a definite advantage in using finite rate of innovation.

The Kolmogorov complexity of these class of signals can be a obtained from the length of the program needed to reconstruct the sequence representing the signal in each period T. The finite rate of innovations of $x(t)$ is defined in terms of the number of degrees of freedom in the given period T. The Kolmogorov complexity of the program which reconstructs the signal depends on the number of degrees of freedom (c_k and t_k in this case). For a signal which can be represented as a finite sum of weighted impulses the Kolmogorov complexity is also finite (here we are considering only the complexity of reconstruction) once we know both c_k and t_k.

If the given signal is an infinite weighted sum of Diracs the finite rate of innovations depend on the choice of T. Therefore the Kolmogorov complexity also depends on T. The number degrees of freedom is a non decreasing function T and therefore the finite rate of innovation and the Kolmogorov complexity of the reconstruction. The method of choosing T for a signal which can be represented as a countably infinite sum of weighted delta functions needs to be investigated.

VI. Discussion and Open Problems

In this work we have provided a Kolmogorov complexity based characterization of signals with finite rate of innovation. The characterization is that such signals are algorithmically compressible. Where, we relate the parametric representation of such signals to their string equivalent, that can be generated by a $K_{\mathcal{U}}(x)$ length program. We feel this characterization is useful to understand farther the algorithmic complexity of sufficient representation, and sampling requirements for signals with finite rate of innovation.

References

[CL78] T. M. Cover and S. K. Leung. Some equivalences between shannon entropy and kolmogorov complexity. *IEEE Trans. Inform. Theory*, IT-24:331–338, 1978.

[CSV84] Ashok K. Chandra, Larry J. Stockmeyer, and Uzi Vishkin. Constant depth reducibility. *SIAM J. Comput.*, 13(2):423–439, 1984.

[CT91] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. John Wiley and Sons, Inc., 1991.

[GV] Peter Grunwald and Paul Vitanyi. Shannon information and kolmogorov complexity. *submitted to IEEE Transactions on Information Theory*.

[HR67] Jr. Hartley Rogers. *Theory of Recursive Functions and Effective Computability*. McGraw-Hill, New York, NY, 1967.

[IMV03] Julius Kusuma Irena Maravic and Martin Vtterli. Low sampling rate uwb channel characterisation and synchronization. *Journal of Communications and Networks, Special issue on ultra-wideband systems*, 5(4):319–327, 2003.

[KRV02] J. Kusuma, A. Ridolfi, and M. Vetterli. Sampling of communication systems with bandwidth expansion. In *In Proc. of the IEEE Conference on Communication*, pages 1601–1605, April 2002.

[LD04] Yue Lu and Minh N. Do. A geometrical approach to sampling sigals with finite rate of innovation. In *ICASSP, 2004 IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2004.

[RMK02] A. Ridolfi, I. Maravic, and J. Kusuma. Sampling signals with finite rate of innovations: Noisy case. In *In Proc. ICASSP*, 2002.

[VMB02] M. Vetterli, P. Marziliano, and T. Blu. Sampling signals with finite rate of innovation. *IEEE Transactions on Signal Processing*, 50(6):1417–1428, June 2002.