MRA OF PROCESSES SYNTHESIZED BY DIFFERINTEGRATION

Onkar J. Dabeer and Uday B. Desai

SPANN Lab.,
Electrical Engineering Department,
IIT Bombay,
Mumbai 400 076, India.
{onkar, ubdesai}@ee.iitb.ernet.in

ABSTRACT

In this paper a definition of multiresoltuion analysis (MRA) of Gaussian processes is proposed. The problem, in a natural way, reduces to the MRA of the associated reproducing kernel Hilbert space. We then show that for processes synthesized from Gaussian white process by fractional integration of order $\alpha \geq 1$, this definition is applicable. The MRA results in an orthogonal expansion of these processes. The region of interest is the positive real line. Using this representation then a decomposition of a wider class of Gaussian processes is given. This representation is multiscale in two ways : firstly, the Gaussian process is split into various component processes characterized by the smoothness of their sample paths and secondly, each of these component processes has a MRA as defined in this paper.

1. INTRODUCTION

MRA of $L^2(\mathcal{R})$ defined by Mallat [8] and the associated wavelet transform have found many applications in the processing of deterministic signals [14]. The wavelet transform has also been used for the analysis of stochastic processes. Wornell [16] has suggested a K-L-like expansion using wavelets to approximate $1/f$ processes. Because of the self-similarity property of fractional Brownian motions, Flandrin [6] has suggested the use of wavelets for their study. Tewfik and Kim [15] have studied the correlation structure of the wavelet coefficients of fractional Brownian motion. Dijkermann and Majumdar [5] have discussed the relation between wavelet transforms of stochastic processes with paths in L^2 and multiresolution models on trees suggested by Basseville *et al* [2]. Krim and Pesquet [7] have suggested a discrete time multiscale framework using which they study processes with stationary increments. Using generalized wavelet packet analysis they have suggested a class of processes with nonstationary increments. The main aim of most of the above mentioned work has been to study and model the wavelet coefficients of stochastic processes. On the other hand the main aim in Benassi and Jaffard [3] is to obtain orthogonal decompositions of Gaussian processes. They give a multiscale, orthogonal decomposition for a large class of Gaussian processes that includes all Markov processes and fractional Brownian motions. This paper uses their methodology to obtain orthonormal bases for the reproducing kernel Hilbert space (RKHS) associated with a Gaussian process.

In this paper a definition of MRA of Gaussian processes is given in Section 3.1 by taking into consideration various properties that are desired from any multiresoltuion scheme. Section 3.2 shows how this definition can be applied to a class of Gaussian processes on $[0,\infty)$ obtained by fractional integration of Gaussian white process. In Section 3.3 we give a decomposition of a larger class of Gaussian processes. Finally, in Section 4, various issues related to the work in this paper and the scope for future research are discussed. To begin with some basic concepts required in this paper are summarized in Section 2.

2. BASIC CONCEPTS

2.1. RKHS associated with a Gaussian process

A Gaussian random process is defined as a collection of random variables $\{X(t) : X(t) \in L^2(\Omega, \mathcal{F}, P), t \in T\}$, where T is some indexing set, $(\Omega, \mathcal{F}, \mathcal{P})$ is a probability space and $L^2(\Omega, \mathcal{F}, \mathcal{P})$ is the Hilbert space of zero mean, finite variance Gaussian random variables. The inner product on this space is defined by, $< x, y >= E(xy)$, where E denotes the expectation. The Hilbert space of random variables generated by this process, $\mathcal{H}_X$, is the closure in $L^2(\Omega, \mathcal{F}, \mathcal{P})$ of the linear span of $X(t), t \in T$. Each random variable z in $\mathcal{H}_X$ is associated with a function f_z on T defined as, $f_z(t) = E(zX(t))$. The inner product between two functions is defined as, $< f_u, f_v >= E(uv)$. With this inner product, these functions constitute a Hilbert space, H_X, that is isomorphic and isometric to $\mathcal{H}_X$. Let $R(t, s)$ be the covariance kernel of the process $X(.)$. It follows that,

$$R(., t) \in H_X \qquad \forall \qquad t \in T \tag{1}$$

$$\langle f_z(.), R(., t)\rangle \quad = E(zX(t)) \quad = f_z(t) \tag{2}$$

These two properties imply that H_X is a reproducing kernel Hilbert space with reproducing kernel $R(t, s)$ (Parzen [11]). Let $\{\phi_\lambda(.)\}$ be an orthonormal basis of H_X. Using (2) and the Parsevals identity it can be shown that $R(t, s) = \sum_\lambda \phi_\lambda(t)\phi_\lambda(s)$ and the process can be synthesized as,

$$X(t) = \sum_\lambda c_\lambda \phi_\lambda(t), \tag{3}$$

where the random variables, c_λ's are s.t $E(c_\lambda c_{\lambda'}) = \delta_{\lambda\lambda'}$. For a detailed discussion of RKHS associated with Gaussian processes the reader is referred to Parzen [11]. A detailed discussion on RKHS can be found in Aronszajn [1].

2.2. Processes characterized by fractional differintegration

2.2.1. Fractional Calculus

Fractional calculus deals with the definition and study of derivatives and integrals of orders that are arbitrary real numbers. There are many ways in which differintegration (following the nomenclature in Oldham and Spanier [10]) can be defined. In this paper we shall have occasion to use the Riemann-Liouville definition for $q < 0$,

$$\left[\frac{d^q f}{dx^q}\right]_{R-L} = \int_0^x (x \Leftrightarrow y)^{-q-1} f(y) dy \tag{4}$$

For $q \geq 0$, following Oldham and Spanier [10], the definition is given by,

$$\frac{d^q f}{dx^q} = \frac{d^n}{dx^n}\left[\frac{d^{q-n} f}{dx^{q-n}}\right]_{R-L} \tag{5}$$

where n is any integer greater than q. A detailed study of fractional calculus is given in Oldham and Spanier [10]. Here only a few properties that are required in this paper are mentioned, namely;

1. The operator $\frac{d^q}{dx^q}$ (henceforth denoted as d^q) is linear for all real q.
2. Scaling : $d^q f(ax) = a^q (d^q f)(ax), a > 0$.
3. A function is said to be differintegrable if it can be written in the form, $f(x) = x^p \sum_{j=0}^{\infty} a_j x^{j/n}$, where $a_0 \neq 0, p > \Leftrightarrow 1$ and n is a integer greater than 0. The following lemma regarding term by term differintegration holds true,

   Lemma 1 *Let $\sum_j f_j$ be a series of differintegrable functions. For $q \leq 0$ term by term differintegration is possible if this series is uniformly convergent over the region of interest and the differintegrated series is also uniformly convergent. For $q > 0$ term by term differintegration is possible if $\sum_j f_j^{(q)}$ and $\sum_j f_j$ are both uniformly convergent in the region of interest.*

4. If $d^q f = g$, then $f(x) = d^{-q} g + c_1 x^{q-1} + ... + c_m x^{q-m}$, where $0 < q \leq m < q + 1$ and $m = 0$ for $q \leq 0$.

2.2.2. Processes synthesized by differintegration

On the positive half line consider the processes that are obtained by the differintegration of order $\Leftrightarrow\alpha$, $\alpha > 0$, of Gaussian white process i.e;

$$X(t) = \int_0^t (t \Leftrightarrow y)^{\alpha-1} dB(y) \tag{6}$$

where $B(\cdot)$ is the Brownian motion and the integral is understood in the mean squared sense. It can be easily seen that the RKHS associated with this process is characterized by the covariance kernel,

$$R(t,s) = \int_0^{\infty} k^{\alpha}(t,y) k^{\alpha}(s,y) dy \tag{7}$$

where $k^{\alpha}(t,y) = (t \Leftrightarrow y)^{\alpha-1}$ for $y \leq t$ and is 0 otherwise. It should be noted that $k^{\alpha}(t,.)$ belongs to $L^2[0,\infty)$ for all $t \geq 0$ and hence, by the Cauchy-Schwartz inequality, the covariance kernel is finite for all $t, s \in [0,\infty)$.

Consider the space defined by, $H = \{f : d^{\alpha} f \in L^2[0,\infty),$ $f^{(\alpha-i)}(0) = 0, i = 0, 1, ..., m\}$, where m is the smallest integer greater than or equal to α. Let H^{α} denote the space obtained by the completion of H under the norm induced by the inner product $\langle f, g\rangle_{H^{\alpha}} = \int_0^{\infty} d^{\alpha} f \, d^{\alpha} g$. That this is a valid inner product can be checked by taking into account the zero initial conditions. The covariance kernel in (7) is also the reproducing kernel of the Hilbert space H^{α}. For proving this, (1) and (2) have to be verified. Since the reproducing kernel completely characterizes a RKHS (Theorem 5b, Parzen [11]), the Hilbert space H^{α} is the RKHS of the process defined in (6). For an orthogonal decomposition of $X(t)$ we want to find orthonormal bases of H^{α}. One way to obtain orthonormal bases is to note that the Hilbert space H^{α} can be considered to be the image of $L^2[0,\infty)$ under the operator $d^{-\alpha}$. The operator $d^{-\alpha}$ preserves the inner product, is continuous and is invertible because of the zero initial conditions (see Property 4 of Section 2.1.1). Hence if $\{\psi_{\lambda}\}$ is an orthonormal basis of $L^2[0,\infty)$, then $\{d^{-\alpha}\psi_{\lambda}\}$ is an orthonormal basis of H^{α}. This way of constructing orthonormal bases for H^{α} will be used in Section 3 to obtain a MRA of the processes synthesized by (6).

2.3. MRA of $L^2[0,\infty)$

A MRA of $L^2[0,\infty)$ can be obtained from a MRA of $L^2(\mathcal{R})$ using a scaling function and a wavelet with compact support [4]. The main idea is to restrict the functions in the multiresolution spaces of $L^2(\mathcal{R})$ to $[0,\infty)$. The translates of $\phi(.)$ and $\psi(.)$ that are unaffected by the restriction (interior functions) are kept as they are and those truncated by the restriction are modified to obtain an equivalent set of edge functions. For a detailed study of the methods used to obtain the edge functions, the reader is referred to Cohen *et al* [4]. We only mention that two scale relations can be derived for this case which characterize a pair of high-pass and low-pass filters corresponding to the interior functions and a pair of filters for each of the edge functions. In this paper the distinction in notation between the edge and interior functions is understood implicitly and is dropped.

3. MRA OF SOME GAUSSIAN PROCESSES

3.1. Definition of MRA

While defining a MRA for a stochastic process one wishes to capture the following ideas,

1. A concept like the MRA of $L^2(\mathcal{R})$ applicable to the collection of sample paths that captures the notion of signal representation at various levels of detail.
2. The random process at a finer level must be expressible as a sum of two independent processes, one of them being a process at the coarser level.
3. The MRA should result in a decomposition of the process in which the coefficients involved are uncorrelated.
4. The MRA should facilitate local study of the process.
5. Given any realization of the process it should be possible to efficiently compute the values attained by the underlying random coefficients.

From (3) it follows that all the functions in the RKHS of the process are sample paths but not all sample paths belong to the RKHS. Thus, it is essential that we have a MRA of the RKHS which should be extendable to the sample paths not included in the RKHS. However since the coarseness of the signal is characterized by the functions in the linear combination and not by the coefficients we shall restrict ourselves a MRA of the RKHS.

Definition 1 *The MRA of a process with RKHS H_X is defined as, (A) A sequence of closed subspaces $...V_{-1} \subset V_0 \subset V_1 \subset ...$, such that : (1) $\overline{\bigcup_{j=-\infty}^{\infty} V_j} = H_X$; (2) $\bigcap_j V_j = \{0\}$; (3) $f(x) \in V_j \Leftrightarrow f(2x) \in V_{j+1}$; (4) $f(x) \in V_0 \Leftrightarrow f(x \Leftrightarrow k) \in V_0 \; \forall \; k \geq 0$;*
(B) A computation strategy that can evaluate the random coefficients in the resulting expansion of the process w.p. 1.

Part (B) of the definition of MRA is new. We believe this to be an important component for the definition of MRA for stochastic processes. The motivation for the condition (B) will become clear in the next section. It will now be shown that MRA defined in this way is possible for processes discussed in Section 2.2.2. Note that for these processes $H_X = H^\alpha$.

3.2. MRA of H_X

Let $\{V'_j\}$ be a sequence of closed subspaces that constitutes a MRA of $L^2[0,\infty)$. Let W'_j denote the orthogonal complement of V'_j in V'_{j+1} and let ϕ and ψ be the scaling function and the wavelet respectively. Then the following theorem holds,

Theorem 1 *Let $V_j = d^{-\alpha} V'_j$. Then the sequence of closed subspaces $\{V_j\}$ satisfies part (A) of the definition of MRA.*

Proof : Property (A.1) : $V_j \subset H_X \; \forall \; j$. Hence $\bigcup_j V_j \subset H_X$. Let $f \in \bigcup_j V_j$. Let $N_f(\epsilon)$ be a neighbourhood of f. Now, $f = d^{-\alpha} g$ for some $g \in V'_k$, for some k. By the continuity of $d^{-\alpha}$, it follows that there is a $\delta > 0$ such that, $d^{-\alpha} N_g(\delta) \subset N_f(\epsilon)$. But $\bigcup_j V'_j$ is dense in $L^2[0,\infty)$. Hence there is a $g_1 \in L^2[0,\infty) \bigcap N_g(\delta)$. Hence, $f_1 = d^{-\alpha} g_1 \in N_f(\epsilon)$. But $d^{-\alpha} L^2[0,\infty)$ is dense in H_X. Hence there is a $f_2 \in H_X$ such that $f_2 \in N_{f_1}(\epsilon)$. It follows that $f_2 \in N_f(2\epsilon)$. This proves property (A.1).
Property (A.2): $f \in \bigcap_j V_j \Rightarrow f = d^{-\alpha} g$, where $g \in \bigcap_j V'_j$. But this implies that $g = 0$. Hence $f = 0$.
Property (A.3): If $f(x) \in V'_j$ then $f(2x) \in V'_{j+1}$. Using this, the desired property now follows from the definition of V_j and Property 2 of Section 2.1.1.
Property (A.4): This can be established easily by a change of variable in the definition of $d^{-\alpha}$. □

Let $W_j = d^{-\alpha} W'_j$. The bases for V_j and W_j are obtained from the corresponding bases for V'_j and W'_j via the operator $d^{-\alpha}$. Also due to property (A.3) and (A.4), except for the edge functions, the other basis functions are translates and dilates of a single function. Due to the linearity of $d^{-\alpha}$ the two scale relations are also the same. Since $\{d^{-\alpha}\psi_{j,k}\}$ is an orthonormal basis for H_X, the process has an orthogonal decomposition of the form,

$$X(t) = \sum_{j=-\infty}^{\infty} \sum_{k=0}^{\infty} c_{j,k} (d^{-\alpha}\psi_{j,k})(t) \tag{8}$$

where the distinction in notation between the edge and the interior functions is dropped and the random coefficients are defined by, $c_{j,k} = \int_0^\infty \psi_{j,k}(y) dB(y)$ and are *i.i.d*, $N(0,1)$.

What is the motivation for condition (B)? Consider the case of Brownian motion ($\alpha = 1$). Given a sample path, to find the coefficient values it is required that the inner product of the sample path with the basis functions be evaluated. But the sample path may not belong to the RKHS, and so, such a method does not have any meaning. In fact, for the Brownian motion, the evaluation of the inner product involves the first derivative of the Brownian path but the Brownian paths can be shown to be differentiable nowhere w. p. 1. (Chapter 1, McKean [9]). Hence, one has to demonstrate how the coefficients can be evaluated for sample paths that are not in the RKHS. Consider the case of $\alpha \geq 1$. Choose an integer n such that $n \Leftrightarrow 1 < \alpha \leq n$. Let ϕ be a scaling function with compact support and ψ be the corresponding wavelet with compact support. By restricting the MRA to a coarsest scale, without loss of generality the 0 scale, the process can be written as,

$$X(t) = \sum_k c'_{0,k} (d^{-\alpha}\phi_{0,k})(t) + \sum_{j=0}^{\infty} \sum_k c_{j,k} (d^{-\alpha}\psi_{j,k})(t) \tag{9}$$

Given a sample path $X(\omega, t)$ we wish to evaluate the values taken by the random coefficients. Let the sample path belong to the RKHS. Then, $c_{j,k}(\omega) = \langle X(\omega,.), d^{-\alpha}\psi_{j,k}(.)\rangle_{H^\alpha}$, and hence,

$$c_{j,k}(\omega) = \int_0^\infty X^{(\alpha)}(\omega,t)\psi_{j,k}(t)dt \tag{10}$$

$$= (\Leftrightarrow 1)^n \int_a^b X^{(\alpha-n)}(\omega,t)\psi_{j,k}^{(n)}(t)dt \tag{11}$$

where in going from (10) to (11) integration by parts has been used. In addition, the fact that $\psi_{j,k}(t)$ has support in some interval $[a,b]$ has been used. More conditions can be imposed on ψ so that formula (11) can be used to evaluate the coefficients even when the sample path is not in the RKHS. The following theorem holds,

Theorem 2 *Let $\alpha \geq 1$ and integer n such that $n \Leftrightarrow 1 < \alpha \leq n$. If the wavelet ψ and the scaling function ϕ are such that,*

- *They have compact support,*
- *they are differintegrable,*
- *their n^{th} derivatives are square integrable,*

then the coefficients in (9) can be evaluated by (11) w.p 1.

We only provide a sketch of the proof here. First the theorem is proved for $\alpha = 1$ and then it is extended to $\alpha > 1$ using Property 3 of Section 2.2.1. The only important step in the proof is the exchange of the order of summation and integration. This is justified by showing that under the assumptions made, the series (8) converges uniformly w.p. 1 on any finite interval. The proof of this fact is a modification of the proof of continuity of Brownian sample paths given in Chapter 1, McKean [9]. The details of this proof are given in [13]. The condition of differintegrability on ϕ and ψ is not very restrictive, as on a compact interval, continuous functions can be uniformly approximated by polynomials and hence most choices of ϕ and ψ will satisfy this condition.

Once the coefficients are evaluated at one scale then the coefficients at the coarser levels can be obtained by discrete filters characterized by the two scale relations. Because of the linearity of the operator $d^{-\alpha}$ the two scale relations are the same as that for the MRA of $L^2[0,\infty)$. The computation strategy for $\alpha \geq 1$ is thus completely specified.

4. DECOMPOSITION OF A WIDER CLASS OF PROCESSES

Theorem 3 *Let $k(x) = \sum_i a_i x^{\alpha_i - 1}$, $\alpha_i > 0$ and a_i are real, be such that the series converges uniformly to $k(x)$ in any finite* *interval. Then the process,* $X(t) = \int_0^t k(t \Leftrightarrow y)dB(y)$ *has a decomposition of the form,*

$$X(t) = \sum_i a_i \sum_{j,k} c^i_{j,k}(d^{-\alpha_i}\psi_{j,k})(t) \qquad (12)$$

where ψ *is a wavelet and the random coefficients* $c^i_{j,k}$ *are such that,* $E(c^i_{j,k}c^n_{l,m}) = \delta_{j,l}\delta_{k,m}$ *and* $c^i_{j,k} = c^n_{j,k}$ *w.p. 1.*

Proof: Because of the assumption of uniform convergence, after substituting for $k(t \Leftrightarrow y)$ in the definition of $X(t)$, the order of integration and summation can be interchanged. Thus the process $X(t)$ is a linear combination of processes of the kind (6). For these processes a decomposition like (8) is possible and (12) follows. The properties of the random coefficients are a direct consequence of their definition.

The multiscale decomposition (12) is characterized by two parameters. One of the parameters is the smoothness (by this we refer to the smoothness of sample paths) level characterized by α_i. Further at each smoothness level a multiresolution structure exists.

5. DISCUSSION

In this paper a definition of MRA of Gaussian processes was motivated and then it was shown that the definition is applicable to a class of nonstationary processes obtained from Gaussian white process by fractional integration of order $\alpha \geq 1$. For the case of $\alpha < 1$ we have proved only part (A) of the definition. Thus for $\alpha < 1$ the processes can be synthesized using (8) or (9) but it is not clear how the coefficients can be evaluated w. p. 1. The decomposition of these processes leads to a decomposition of a wide class of processes mentioned in Theorem 3. The decomposition (12) captures the notion of multiresolution in two ways. The process is split into component processes of different smoothness and the process at each level of smoothness has a MRA as defined in Section 3.1. The class of processes that can be studied in this way is indeed very wide. In particular, Gaussian processes synthesized from Gaussian white process by the inverse systems of constant coefficient differential operators are included in this class. This can be shown by considering the power series expansion of the exponential function. The differential operators are of particular interest because of the fact that Markov processes are characterized by local operators (Pitt [12]). Again, we have not dealt with the issue of evaluating the coefficients in (12). This is a direction for future work.

Some issues regarding the representation (8) and (12) have not been mentioned in this paper due to lack of space. One of the critical issues, from the point of view of local analysis, is the support and decay of the basis functions for the RKHS. By a proper choice of the original wavelet it is possible to do local analysis. Another important issue is that of approximation. It is not very difficult to study the variation with scale of the L^∞ and L^2 norms of the basis functions along the smoothness level as well as along the resolution level. These calculations can be used to obtain approximations to the process. Finally which processes have a decomposition of the form (8) ? A necessary condition for such a decomposition is that $R(2t, 2s) = constant \times R(t, s)$. For example a process synthesized from Gaussian white noise from the *inverse* of the operator $D = \frac{d}{dx} + a$, $a \neq 0$ has a covariance kernel, $R(t,s) = \frac{e^{a(t+s)}}{2a}(1 \Leftrightarrow e^{-2a(t \wedge s)})$ which does not satisfy the above mentioned property. The reason for this is that the operator D does not commute with the dilation operation up to a scale factor and hence condition (A.3) is not satisfied. Hence the definition of MRA has to be modified if more processes are to be analyzed in a MRA framework and (12) indicates a natural way to do so.

6. REFERENCES

[1] N. Aronszajn, "Theory of reproducing kernels", *Trans. of American Mathematical Society*, pp 337-404, May 1950.

[2] M. Basseville, A. Benveniste, K. Chou, S. A. Golden, R. Nikoukhah and A. Wilsky, "Modelling and estimation of multiresolution stochastic processes", *IEEE Trans. Information Theory*, vol. 38, no. 2, pp 766-784, March 1992.

[3] A. Bennasi and S. Jaffard,"Wavelet decomposition of one and several dimensional Gaussian processes", *Recent Advances in Wavelets*, L. L. Schumaker and G.Webb (eds.), Academic Press Inc., pp 119-154, 1994.

[4] A. Cohen, I. Daubechies and P. Vial, " Wavelets on the interval and Fast Wavelet Transforms", *Applied and Computational Harmonic analysis*, vol. 1, pp 54-81, 1993.

[5] R. W. Dijkermann and R. R. Majumdar, "Wavelet transform of stochastic processes and multiresolution stochastic models", *IEEE Trans. Signal processing*, vol. 42, no. 7, pp 1640-1652, July 1994.

[6] P. Flandrin, "Wavelet Analysis and synthesis of fractional Brownian motion", *IEEE Trans. Information Theory*, vol. 38, no. 2, pp 910-917, March 1992.

[7] H. Krim and J. C. Pesquet, "Multiresolution analysis of a class of nonstationary processes", *IEEE Trans. Information Theory*, vol. 41, no. 4, pp 1010-1020, July 95.

[8] S. G. Mallat, "Multiresolution approximations and wavelet orthonormal bases of $L^2(\mathcal{R})$", *Trans. of American Mathematical Society*, vol. 35, no. 1, pp 69-87, Sept. 1989.

[9] H. P. McKean, "Stochastic Integrals", Academic press, New York, 1969.

[10] K. B. Oldham and J. Spanier, " The Fractional Calculus ", Academic Press, New York and London, 1974.

[11] E. Parzen, " Statistical inference on time series by Hilbert space methods, I ", Time Series Analysis papers, Holden Day, 1967.

[12] L. D. Pitt, " A Markov property for Gaussian processes with a multi-dimensional parameter", *Archives of Rational Mechanics*, 43, pp 367-391, 1971.

[13] O. J. Dabeer and U. B. Desai, "MRA of stochastic processes", *SPANN Lab. Technical Report*, *SPANN 97:O1*, Electrical Engineering Dept., IIT, Bombay, 1997.

[14] Special issue on Wavelets, *Proceedings of the IEEE*, April 1996.

[15] A. H. Tewfik and M. Kim, "Correlation structure of the discrete wavelet coefficients of fractional Brownian motion", *IEEE Trans. Information Theory*, vol 38, no. 2, pp 904-909, March 1992.

[16] G. W. Wornell, "A Karhunen-Loeve-like expansion for 1/f processes via wavelets", *IEEE Trans. Information Theory*, vol 36, pp 859-861, July 1990.