Random Wavelet Transforms, Algebraic Geometric Coding, and Their Applications in Signal Compression and De-Noising

Tomasz Bieleck
Li M. Song*
Stephen S. T. Yau*
Department of Mathematics, Statistics and Computer Science
The University of Illinois at Chicago
Chicago, IL 60607-7045

Man Kam Kwong†
Mathematics and Computer Science Division
Argonne National Laboratory
Argonne, IL 60439-4844
E-mail: kwong@mcs.anl.gov

Abstract

The concepts of random wavelet transforms and discrete random wavelet transforms are introduced. It is shown that these transforms can lead to simultaneous compression and de-noising of signals that have been corrupted with fractional noises. Potential applications of algebraic geometric coding theory to encode the ensuing data are also discussed.

Introduction

In this paper, we first outline the main ideas behind classical wavelet-based algorithms of compression and de-noising of signals. We then introduce the concept of random wavelet transforms and discrete random wavelet transforms, which are more effective than their classical counterparts in handling fractional noises. Finally, we discuss potential applications of the method of algebraic geometric coding of data produced by our new algorithms.

Since this article is not a detailed survey, many aspects of both compression and coding algorithms will not be presented here. For example, we do not discuss thresholding algorithms; we limit discussion of wavelet bases to the orthonormal ones generated by the well-known Daubechies scaling functions (see [13,14] for a discussion of a new method to construct biorthogonal bases); and we do not give a detailed analysis of coding algorithms. Instead, this paper is meant to be an overview of a methodology dealing with a specific class of signal+noise models that we have in mind.

In Section 1, we describe a class of noisy signal models. In particular, we focus on fractional random fields, which occur naturally in many applications. In Section 2, we formulate a signal de-noising problem. In solving this problem, we achieve, in addition, a compression of the signal. In Section 3, we present some ideas from wavelet analysis that motivate our approach. In particular, we lay the ground for the so-called discrete wavelet transform that we introduce in Section 4. Section 4 is pivotal for the paper: it presents the essence of the simultaneous compression/de-noising algorithms. In the last two sections (5 and 6) we present briefly some new ideas from algebraic geometric coding theory which, we believe, can be used beneficially in source and channel coding for image-based information, especially that resulting from our compression/de-noising algorithms. Novelty and strength of the algebraic geometric codes are discussed.

1. Noisy Signal Models

We begin by describing the class of noise processes that we are concerned with. Since we are going to deal with multiparameter signals in general, the corrupting noise will be modeled as a random field parametrized by $t \in \mathbf{R}^d$. The cases $d=1$ and $d=2$ correspond to one- and two-dimensional signals, respectively.

For each $t \in \mathbf{R}^d$, let C_t denote the cuboid determined by t and 0, and let V_{C_t} denote the Euclidean volume of C_t. We suppose that a probability space $(\Omega, \mathcal{F}, P)$ has been defined.

Definition 1.1 We define a fractional Wiener sheet (fWs) on $\mathbf{R}^d$ with parameter $H \in (0,1]$ to be a real-valued Gaussian random field $W^H \equiv (W^H(t))_{t \in \mathbf{R}^d}$ such that

i) $W^H(0) = 0, \quad$ P a.s.,

ii) $EW^H(t) = 0, \quad \forall t \in \mathbf{R}^d$,

iii) $E(W^H(t) - W^H(s))^2 = k_H \cdot |V_{C_t} + V_{C_s} - 2V_{C_t \cap C_s}|^{2H}$, $\forall t, s \in \mathbf{R}^d$, $k_H = \text{const} > 0$,

---

*The Control and Information Laboratory

†Work supported by the Mathematical, Information, and Computational Sciences Division subprogram of the Office of Computational and Technology Research, U.S. Department of Energy, under Contract W-31-109-Eng-38.

where E denotes the expectation w.r.t. the probability function P.

Remark 1.1 A fractional Wiener sheet on $\mathbf{R}^d$ was first constructed and studied in [2].

Remark 1.2 In the case $d = 1$, a fractional Wiener sheet coincides with fractional Brownian motion (fBm) and is denoted by $B^H \equiv (B^H(t))_{t\in\mathbf{R}}$. For the classical definition of fBm, see, for example, [18]. A discussion of various real-life phenomena giving rise to fBm can be found in [12]. The fWs extends the concept of fBm to the multiparameter case and is very useful in modeling multiparameter disturbances whose intensities depend on the size of the perturbed object.

Remark 1.3 Note that in the case $H = \frac{1}{2}$, the fWs is equivalent to an ordinary Brownian motion with a d-dimensional "*time*" parameter.

Let $\dot{W}^H \equiv (\dot{W}^H(t))_{t\in\mathbf{R}^d}$ be a random field consisting of distributional derivatives of the fWs W^H w.r.t. $t \in \mathbf{R}^d$. We shall call such a random field a fractional Wiener noise (fWn). In the case $H = \frac{1}{2}$, $\dot{W}^H$ is the Gaussian white noise on $\mathbf{R}^d$.

The class of noisy signals that we are considering in this paper consists of random fields $(R(t))_{t\in\mathbf{R}^d}$ such that

$$R(t) = S(t) + W^H(t), \quad t \in \mathbf{R}^d \tag{1.1}$$

or

$$R(t) = S(t) + \dot{W}^H(t), t \in \mathbf{R}^d \tag{1.1'}$$

where $S(\cdot) \in L^2(\mathbf{R}^d)$. The function $S(\cdot)$ is called the original signal and $R(\cdot)$ the recorded (or measured, or observed) signal.

2. The Problem of Signal De-noising

Simply put, the abstract problem of signal de-noising amounts to estimating $S(\cdot)$ from the noisy observation $R(\cdot)$. From the standpoint of statistical theory, it is a regression problem and, most typically, a nonparametric one.

In practice, one does not have access to the entire $R(\cdot)$, but only to a finite number of sampled values: either $R(t_i)$ or some average of $R(\cdot)$ in a small neighborhood of t_i for $i = 1, \cdots, I$. The practical problem, therefore, is to estimate $S(\cdot)$ given the data $R(t_i), i = 1, \cdots, I$ and the relation (1.1) or (1.1′).

Let $F = \{f_k, k \in K\}$ be a basis for $L^2(\mathbf{R}^d)$. We can write

$$R(t_i) = \sum_{k\in K} c_k f_k(t_i) + W^H(t_i), \quad i = 1, \cdots, I, \tag{2.1}$$

or

$$R(t_i) = \sum_{k\in K} c_k f_k(t_i) + \dot{W}^H(t_i), \quad i = 1, \cdots, I, \tag{2.1'}$$

where $c_k, k \in K$ are the coefficients of $S(\cdot)$ in the basis F. The estimation problem for $S(\cdot)$ may therefore be reduced to that of the c_k's. Obviously, one cannot hope to get good estimates of all the c_k's when using only a finite number I of observations $R(t_i)$. Therefore the expansions in (2.1) and (2.1′) are truncated to a finite number of coefficients, and one estimates only the coefficients that are retained. It is clear that the success of this method depends, to a large extent, on the accuracy of the approximation of $S(\cdot)$ by the reduced expansion. It turns out that very good results can be obtained if one uses the so-called wavelet basis to carry out the expansions in (2.1) or (2.1′).

3. Wavelet Transform and Random Wavelet Transform

We will discuss, for simplicity, only one-dimensional wavelets. Higher-dimensional wavelets can be treated in a similar way. In her seminal paper [5], Ingrid Daubechies presented the construction of two sequences (p_k) and (q_k) of real numbers that solve the two-scale equations,

$$\phi(x) = \sum_k p_k \phi(2x - k) \tag{3.1}$$

$$\psi(x) = \sum_k q_k \psi(2x - k) \tag{3.2}$$

for all $x \in \mathbf{R}^{\mathbf{d}}$.

She proved the existence of compactly supported solutions in $L^2(\mathbf{R})$. For each solution, the functions $\phi(\cdot)$ and $\psi(\cdot)$ are called the scaling function and mother wavelet, respectively. The name wavelet for $\psi(\cdot)$ is motivated by the following property:

$$\int_{\mathbf{R}} \psi(x)dx = 0. \tag{3.3}$$

A very important role in wavelet theory is played by the translated and dilated version of the basic functions $\phi(\cdot)$ and $\psi(\cdot)$,

$$\phi_{j,k}(x) := 2^{j/2}\phi(2^j x - k) \tag{3.4}$$

$$\psi_{j,k}(x) := 2^{j/2}\psi(2^j x - k) \tag{3.5}$$

for all $x \in \mathbf{R}, j, k \in \mathbf{Z}$.

Let us denote, for $\mathcal{J} \in \mathbf{Z}$,

$$\Phi_{\mathcal{J}} = \{\phi_{\mathcal{J},k}(\cdot), k \in \mathbf{Z}\},$$

and

$$\Psi_{\mathcal{J}} = \{\psi_{j,k}(\cdot), j \geq \mathcal{J}, k \in \mathbf{Z}\}.$$

Then, as shown in [5], $(\Phi_{\mathcal{J}}, \Psi_{\mathcal{J}})$ is an orthonormal basis for $L^2(\mathbf{R})$ for each $\mathcal{J} \in \mathbf{Z}$, under appropriate choice of the coefficients (p_k) and (q_k). Such a choice of (p_k) and (q_k) will be assumed from now on. In particular, $\Psi_{-\infty} = \{\psi_{j,k}, j, k \in \mathbf{Z}\}$ is an orthonormal wavelet basis for $L^2(\mathbf{R})$.

For each locally square integrable function f, we define the scaling coefficients $(a_{j,k})$ and wavelet coefficients $(b_{j,k})$ as the respective integral transforms

$$a_{j,k} := \int_{\mathbf{R}} f(x)\phi_{j,k}(x)dx, \tag{3.6}$$

and

$$b_{j,k} := \int_{\mathbf{R}} f(x)\psi_{j,k}(x)dx \tag{3.7}$$

for all $j, k \in \mathbf{Z}$. The right hand side of (3.7) is called the wavelet transform of $f(\cdot)$. If $f \in L^2(\mathbf{R})$ then we have

$$f(x) = \sum_{k=-\infty}^{\infty} a_{\mathcal{J},k}\phi_{\mathcal{J},k}(x) + \sum_{j=\mathcal{J}}^{\infty}\sum_{k=-\infty}^{\infty} b_{j,k}\psi_{j,k}(x) \tag{3.8}$$

and

$$f(x) = \sum_{k=-\infty}^{\infty}\sum_{j=-\infty}^{\infty} b_{j,k}\psi_{j,k}(x), \tag{3.9}$$

where the expansions are in the $L^2(\mathbf{R})$ sense. The following notation will be useful.

$$V_j := closure_{L^2(\mathbf{R})}\{\phi_{j,k}, k \in \mathbf{Z}\}, \quad V_{-\infty} := \phi, \tag{3.10}$$

$$W_j := closure_{L^2(\mathbf{R})}\{\psi_{j,k}, k \in \mathbf{Z}\}.$$

Then we have (see [5] or [4])

$$V_{j+1} = V_j \bigoplus W_j. \tag{3.11}$$

Therefore, if we denote by f^j the projection of f on V_j, and by g^j the projection of f on W_j, we see that for each $f \in L^2(\mathbf{R})$ and all $M \geq N$ we have

$$f^M = f^N + g^N + g^{N+1} + \cdots + g^{M-1}. \tag{3.12}$$

The projection f^M is thus obtained by adding various levels of "detail", encoded in g^j's, to the "blurred" version f^N of f^M. In particular, we have

$$f = f^N + \sum_{j=N}^{\infty} g^j, \tag{3.13}$$

which is the same as (3.8), and

$$f = \sum_{j=-\infty}^{\infty} g^j, \tag{3.14}$$

which is the same as (3.9).

Let us denote by $(c_k^j, k \in \mathbf{Z})$ and $(d_k^j, k \in \mathbf{Z})$ the coefficients of f in the basis of V^j and W^j, respectively, for each $j \in \mathbf{Z}$. That is,

$$f^j = \sum_k c_k^j \phi_{j,k}, \tag{3.15}$$

$$g^j = \sum_k d_k^j \psi_{j,k} \tag{3.16}$$

for each $j \in \mathbf{Z}$. There exists a unique pair of numerical sequences $\{(\tilde{p}_k), (\tilde{q}_k)\}$ so that one can obtain the coefficients $(c_k^j, k \in \mathbf{Z})$, $(d_k^j, k \in \mathbf{Z})$, $j = M-1, M-2, \cdots, N$, from the coefficients $(c_k^M, k \in \mathbf{Z})$, by means of the following "decomposition algorithm":

$$c_k^{j-1} = \sum_l \tilde{p}_{l-2k} c_l^j \tag{3.17}$$

$$d_k^{j-1} = \sum_l \tilde{q}_{l-2k} d_l^j, \qquad j = M, M-1, \cdots, N. \tag{3.18}$$

It is also possible to "reconstruct" the c_k^M's from the sequences $(c_k^j, k \in \mathbf{Z})$ and $(d_k^j, k \in \mathbf{Z})$, $j = N, \cdots, M-1$, by means of the "reconstruction algorithm":

$$c_k^j = \sum_l (p_{k-2l} c_l^{j-1} + q_{k-2l} d_l^{j-1}), \qquad j = N+1, \cdots, M, \tag{3.19}$$

where p_k, q_k are the original Daubechies coefficients.

The above algorithms are the Mallat pyramidal algorithms ([17]).

It is important to note that the coefficients $(c_k^N, k \in \mathbf{Z})$, $(d_k^j, k \in \mathbf{Z})$, $j = N, N+1, \cdots, M-1$, can be obtained from $(c_k^M, k \in \mathbf{Z})$ in one shot. Let us denote by $(r_l^{M,N,k}, l \in \mathbf{Z})$ and $(s_l^{M,j,k}, l \in \mathbf{Z})$ the coefficients of $\phi_{N,k}$ and $\phi_{j,k}$, $j = N, N+1, \cdots, M-1$, respectively, in the basis of V_M. That is,

$$\phi_{N,k} = \sum_l r_l^{M,N,k}\phi_{M,l} \tag{3.20}$$

$$\psi_{j,k} = \sum_l s_l^{M,j,k}\psi_{M,l} \tag{3.21}$$

for $k \in \mathbf{Z}$ and $j = N, N+1, \cdots, M-1$. Then

$$c_k^N = \sum_l r_l^{M,N,k} c_l^M \tag{3.22}$$

$$d_k^j = \sum_l s_l^{M,j,k} c_l^M \tag{3.23}$$

for $k \in \mathbf{Z}$ and $j = N, N+1, \cdots, M-1$. The formulas (3.23) and (3.24) express the simultaneous projections of f^M on the spaces $V_N, W_N, W_{N+1}, \cdots, W_{M-1}$. We can write (3.22) and (3.23) in matrix form as

$$\begin{bmatrix} \mathbf{c}^N \\ \mathbf{d}^N \\ \mathbf{d}^{N+1} \\ \vdots \\ \mathbf{d}^{M-1} \end{bmatrix} = \begin{bmatrix} \mathbf{r}^{M,N} \\ \mathbf{s}^{M,N} \\ \mathbf{s}^{M,N+1} \\ \vdots \\ \mathbf{s}^{M,M-1} \end{bmatrix} \mathbf{c}^M, \tag{3.24}$$

where

$$\mathbf{c}^i = [c_l^i]_{l\in\mathbf{Z}}$$

$$\mathbf{d}^j = [d_l^j]_{l\in\mathbf{Z}}$$

$$\mathbf{r}^{M,N} = [r_l^{M,N,k}]_{l,k\in\mathbf{Z}}$$

$$\mathbf{s}^{M,j} = [s_l^{M,j,k}]_{l,k\in\mathbf{Z}}$$

for $i = M, N$, $j = N, N+1, \cdots, M-1$. If we denote the left hand side of (3.24) by $\mathbf{w}^{M,N}$, and the matrix on the right-hand side by $\mathbf{W}^{M,N}$, we can write it as

$$\mathbf{w}^{M,N} = \mathbf{W}^{M,N} \cdot \mathbf{c}^M. \tag{3.25}$$

This is the key relation for the discrete wavelet transform to be introduced in Section 4.

In the remaining part of this section, we shall briefly describe transforms of the fractional Brownian motion B^H. The results given below can be extended to the case of Wiener sheet by using the result of [2].

The scaling and wavelet coefficients of B^H are random sequences given by

$$a_{j,k}^H = \int_{\mathbf{R}} B^H(x)\phi_{j,k}(x)dx, \tag{3.26}$$

and

$$b_{j,k}^H = \int_{\mathbf{R}} B^H(x)\psi_{j,k}(x)dx \tag{3.27}$$

respectively, for $j, k \in \mathbf{Z}$. Using the results of [1], we can also define what we call random scaling and random wavelet coefficients of B^H. We have used stochastic integrals of scaling and wavelet functions with respect to B^M. That is, we define

$$\bar{a}_{j,k}^H = \int_{\mathbf{R}} \phi_{j,k}(x)dB^H(x) \tag{3.28}$$

and

$$\bar{b}_{j,k}^H = \int_{\mathbf{R}} \psi_{j,k}(x)dB^H(x) \tag{3.29}$$

for $j, k \in \mathbf{Z}$. Several properties of the random wavelet coefficients of B^M are demonstrated in [3]. The transform in (3.29) is called the random wavelet transform.

The scaling and wavelet transforms of $\dot{B}^H$ can be defined in terms of (3.29) and (3.30) as

$$\int_{\mathbf{R}} \phi_{j,k}(x)\dot{B}^H(x)dx = \bar{a}_{j,k}^H \tag{3.30}$$

$$\int_{\mathbf{R}} \psi_{j,k}(x)\dot{B}^H(x)dx = \bar{b}_{j,k}^H \tag{3.31}$$

for $j, k \in \mathbf{Z}$.

For $d = 1$, the models (1.1) and (1.1′) can be written as

$$R(t) = S(t) + B^H(t) \tag{3.32}$$

or

$$R(t) = S(t) + \dot{B}^H(t) \tag{3.32′}$$

for $t \in \mathbf{R}$. This implies the following relations for the respective scaling and wavelet coefficients,

$$a_{j,k}^R = a_{j,k}^S + a_{j,k}^M \tag{3.33}$$

$$b_{j,k}^R = b_{j,k}^S + b_{j,k}^M \tag{3.34}$$

or

$$a_{j,k}^R = a_{j,k}^S + \bar{a}_{j,k}^M \tag{3.33′}$$

$$b_{j,k}^R = b_{j,k}^S + \bar{b}_{j,k}^M \tag{3.34′}$$

for $j, k \in \mathbf{Z}$. The above relations are used in the compression and de-noising algorithms given in the next section.

4. Discrete Wavelet Transform and Signal De-Noising

Suppose that the $I = 2^M$ coefficients of the projection f^M of f on V_M, $M > 0$, are known. Using these I coefficients, we can compute approximations to $c_0^0, d_0^0, d_0^1, d_1^1, \cdots, d_0^{M-1}, \cdots, d_{I/2-1}^{M-1}$, the coefficients introduced in the preceding section. We shall use a finite-dimensional counterpart of (3.25). In other words, denoting by $\mathbf{c}_I$ the vector of I known coefficients of f^M and by $\mathbf{W}_I$ the appropriate finite dimensional part of $\mathbf{W}^{M,0}$, we define the vector $\mathbf{w}_I$ of "discrete" wavelet coefficients of f by

$$\mathbf{w}_I := \mathbf{W}_I \mathbf{c}_I. \tag{4.1}$$

The matrix transform in (4.1) is called the discrete wavelet transform. The vector $\mathbf{w}_I$ approximates the coefficient vector $[c_0^0, d_0^0, d_0^1, d_1^1, \cdots, d_0^{M-1}, \cdots, d_{I/2-1}^{M-1}]$.

It is not difficult to see that W_I is an orthogonal matrix. Therefore we have

$$\mathbf{c}_I := \mathbf{W}_I^T \mathbf{w}_I, \tag{4.2}$$

where T denotes the transposition operation. The matrix transform in (4.2) is called the inverse discrete wavelet transform.

An important feature of (4.1) is that, typically, many of the elements of $\mathbf{w}_I$ are near zero and therefore can be discarded for compression and estimation purposes without significantly altering the outcome of the inverse transformation (4.2). This property is the basis for the idea of a simultaneous signal compression and de-noising that we shall briefly discuss below.

Let $\mathbf{c}_I^R, \mathbf{c}_I^S, \mathbf{c}_I^H$, and $\overline{\mathbf{c}}_I^H$ denote the vectors of coefficients obtained by convolving $R(\cdot), S(\cdot)$, $B^H(\cdot)$ and $\dot{B}^H(\cdot)$, respectively, with I elements of the basis of V_M. Only the coefficients c_I^R are observed in practice, in other words, the elements of $\mathbf{c}_I^R$ are approximated by sampled measurements of $R(\cdot)$ at I points.

Now, we can obviously write

$$\mathbf{c}_I^R = \mathbf{c}_I^S + \bar{\mathbf{c}}_I^M \tag{4.3}$$

or

$$\mathbf{c}_I^R = \mathbf{c}_I^S + \bar{\mathbf{c}}_I^M. \tag{4.3′}$$

Applying (4.1) to (4.3) or (4.3′) we get

$$\mathbf{w}_I^R = \mathbf{w}_I^S + \mathbf{w}_I^M \tag{4.4}$$

or

$$\mathbf{w}_I^R = \mathbf{w}_I^S + \bar{\mathbf{w}}_I^M. \tag{4.4$'$}$$

As we indicated earlier, many elements of $\mathbf{w}_I^R$ will be near zero. Using appropriate thresholding, we first set to zero those elements of $\mathbf{w}_I^R$ that are appropriately small. In other words, we compress $\mathbf{w}_I^R$. After that, we estimate elements of $\mathbf{w}_I^S$ by using whatever information is left in the compressed version of $\mathbf{w}_I^R$. There are, of course, numerous estimation procedures possible. The one that we used in our simulations was based on appropriate shrinkage of the elements of the compressed $\mathbf{w}_I^R$. Actually, we have achieved a simultaneous compression and estimation (de-noising) in one shrinkage operation.

Let us denote by $\hat{\mathbf{w}}_I^S$ the estimate of $\mathbf{w}_I^S$. Applying (4.2) to $\hat{\mathbf{w}}_I^S$, we obtain the estimate of $\mathbf{c}_I^S$, as

$$\hat{\mathbf{c}}_I^S = \mathbf{W}_I^T \hat{\mathbf{w}}_I^S. \tag{4.5}$$

Note that (4.5) represents a simultaneous decompression and estimation of the original, unobserved signal $S(\cdot)$. If we denote by ${}^s\hat{c}_0^0, {}^s\hat{d}_0^0, {}^s\hat{d}_0^1, {}^s\hat{d}_1^1, \cdots, {}^s\hat{d}_0^{M-1}, \cdots, {}^s\hat{d}_{I/2-1}^{M-1}$ the elements of $\mathbf{c}_I^S$, then the estimate $\hat{S}(\cdot)$ is obtained by

$$\hat{S}(t) = {}^s\hat{c}_0^0\phi_{0,0}(t) + {}^s\hat{d}_0^0\psi_{0,0}(t) + {}^s\hat{d}_0^1\psi_{1,0}(t) + {}^s\hat{d}_1^1\psi_{1,1}(t) + \cdots + {}^s\hat{d}_0^{M-1}\psi_{M-1,0}(t) + \cdots + {}^s\hat{d}_{I/2-1}^{M-1}\psi_{M-1,I/2-1}(t) \tag{4.6}$$

for $t \in \mathbf{R}$.

5. Digital Communication

The methodology of Sections 1 through 4 produces a vector of data $\hat{\mathbf{w}}_I$ that one may wish to store and/or to transmit. In this and the next section, we present some basic principles of algebraic geometric coding that we intend to use to code the data $\hat{\mathbf{w}}_I$ and transmit them via digital communication channels.

Digital communication offers various advantages and has become increasingly important. Some of its advantages are its flexibility, reliability, and availability of wide-band channels such as optical fibers and satellite. A typical (memoryless) digital communication system has these channels:

$$\textit{digital source} \rightarrow \textit{source encoder}$$
$$\rightarrow \textit{channel encoder} \rightarrow \textit{modulator}$$
$$\Rightarrow \text{Noisy Channel} \Rightarrow$$
$$\textit{detector} \rightarrow \textit{channel decoder}$$
$$\rightarrow \textit{source decoder} \rightarrow \textit{digital sink}$$

where the effect of the detector is a "demodulator." Source coding (here a form of lossless data compression) is a means to remove redundancy. It maps the digital source into some code, where full recovery is possible, seeking to represent the source efficiently in the sense that the average length (bit per symbol) is minimal. Such average length is bounded below by the entropy, the information that the source carries (known as the source coding theorem; see [10]). Low average length can generally be made close to the entropy for the price of decoding complexity.

In the approach presented in this paper, random wavelet representation is the main tool we use in signal compression/de-noising. This compression procedure, accomplished with various thresholding techniques, is an entropy-reduction transformation. It actually appears before the steps depicted in the above diagram, and produces the "digital source." Understanding the statistics of the source (which in our case is the statistical distribution of the wavelet representation of the signals) or the statistics of the wavelet coefficients can lead to significant entropy reduction, thus increasing the compression ratio. This feature reinforces the importance of the noise modeling. The actual source coding involves encoding the coefficients in the random wavelet representation of the signal.

6. Algebraic Geometric Codes

1. Channel Coding.

The objective of channel coding (error control) is to achieve reliability of transmitting information through a noisy channel. Channel coding introduces redundancy in order to control the error as a result of channel noise during transmission. It seems to be the only practical way to achieve reliability and efficiency.

A channel is described by the conditional probability of correct reception The term channel capacity, C, is introduced as the maximal mutual information to measure the capacity of the channel ([10]). A classical theorem of Shannon (the channel coding theorem; compare the source coding theorem in the preceding section) states that if

$$\frac{H(\mathcal{S})}{T_s} \le \frac{C}{T_c},$$

where, T_s is the source rate measured in symbol per T_s seconds, T_c is the channel rate in symbol per T_c seconds, and $H(\mathcal{S})$ is the entropy, then there exists some code with arbitrary small error probability.

Two parameters are apparent to the designer: transmitted signal power and bandwidth. These parameters in turn, through a modulation scheme, determine the ratio of the signal energy per bit E_b to the noise power density N_0. The reduction of this ratio E_b/N_0 can mean lower transmitted power requirement and hardware cost, and thus serves as some measure of performance. Coding is the only practical way to achieve a small error probability Furthermore, careful design and choice of coding scheme can lead to the reduction of E_b/N_0, known as the code gain. This is one factor that prompts us to consider algebraic geometric codes.

2. Algebraic Geometric Codes.

There are two types of code in common practice, block codes and convolutional codes. Among block codes, linear code is the most common for its easy manipulation (because of the extra structure it carries). In this paper, we will not be concerned with convolutional codes. The economics of a linear code is measured by the parameters $[n, k, d]$ (sometimes $[n,k,d]_q$ to signify the presence of the underlying finite field) of length n, dimension k, and minimal distance d. It is a k-dimensional vector subspace sitting inside some standard n-dimensional vector space over a given finite field, say $\mathbf{F}_q$, through an encoding process. Various ingenious techniques (such as concatenation) can improve code performance. Classical constructions (typically cyclic codes) usually depend in some way on the parameters n and k (and d). It has been the case that, as the codeword length n increases, either k or d gains poorly [15]; that is, *dnumber* or *knumber* will tend to 0. In other words, the code cannot compensate satisfactorily for both error-correcting ability and information rate.

Algebraic geometric (AG) codes come to the rescue. They were constructed about fifteen years ago, an interdisciplinary fruit of coding theory and algebraic geometry, and were soon proven to be of theoretic importance. They are the first examples to exceed the Gilbert-Varshamov bound; both *dnumber* and *knumber* are bounded away from 0 when n becomes large (indeed, far better bounds). Such measures are important so that we can check taht the code rate and error-correcting capability do not diminish miserably. The complexity incurred from aspects such as the increase of codeword length (hence bandwidth requirement) can now be analyzed.

3. Construction of Algebraic Geometric Code*.

An AG code is constructed ([16,23]) from a smooth projective curve over a finite field $\mathbf{F}_q$. In one way it is the image of the evaluation map from a certain space of sections $L(D)$ at some n rational points P_i. Here D is a divisor with support disjoint from the divisor $P := P_1 + ... + P_n$. In most practical situations these sections are represented by polynomials, and the evaluation is done at specific points (e.g., on the projective plane). By Weil's conjecture ([9]) for curves there are approximately the scale of $2q$ (rational) points available on the curve. Clearly, such construction provides us with numerous choices to construct linear codes. Although the actual minimal distance is hard to determine because the theorem of Riemann-Roch does not tell us how to compute the dimension of the space of sections associated to a special divisor, we can obtain a comparable quantity $d^* = deg(G) - 2g + 2$ which plays a role in measuring the code's error-correcting ability, where g is the genus of the curve. This is called the designed minimal distance.

To spell out some detail, let the above defined code have parameters $[n, k, d]$. If we assume that $2g - 2 < deg(G) < n$, then $k = deg(G) - g + 1$, and $d > d^*$ ([17]). So for choice of $deg(G)$ at the scale of n, it is clear that both k and d will be of the same scale.

*We treat this topic informally.

4. Decoding Algorithms.

Decoding algorithms are one of the most important factors in channel coding. The study in this area began late last decade; the efforts are summarized in [23]. The basic algorithm is similar to the idea of decoding cyclic codes: (1) write down a parity check matrix; (2) determine an error locator; (3) solve a system of linear equations to determine the error location; and (4) finally evaluate the correct value. But the error locator does not consist exactly of error locations; extra places may be introduced. The point has always been how to solve a system of linear equations (not an arbitrary one) effectively.

Three recent results in the area of decoding AG codes are worth attention. They have close error-correcting capacity (by which we mean the number of errors the code is capable of correcting) and similar complexity, as illustrated below, and are given respectively in Justesen et. al. [11], Feng and Rao [7], and Ehrhard [6].

Ref.	Curve	Capacity	Complexity
[11]	planar	$\frac{d^*i}{2} - \frac{m^2}{8} + \frac{m}{4} - \frac{9}{8}$	$O(n^{\frac{7}{3}})$
[7]	any	$(d^* - 1)/2$	$O(n^3)$
[6]	any	$(d^* - 1)/2$	$O((d^*)^2 n)$

In the above, m is the degree of the plane curve; d^* should be considered as proportional to n. They each were able to make use of an algorithm generalizing some standard ones to the cyclic codes (cases of [11] and [7]), thanks to the underlying algebraic structure. An equivalent description using differentials was the presentation in [7] and [6]. They all have the appeal of implementability.

5. Perspective.

AG codes are easy to construct, have rates and error-correcting capacity, and algorithms for decoding are fast. In [24] the authors constructed a code, for channels with or without memory, with a rate arbitrarily close to the channel capacity, and with the remarkable property (an example of the identification coding theorem of

Ahlswede and Dueck) that the transmission can achieve double exponential rate. Two constructions were given. One was a triple-layered concatenated code; the other used algebraic geometric codes that were equipped with extraordinary bounds. It is also known [21] that algebraic geometric codes could be used to rewrite a nonlinear code into a linear code.

We conclude that AG codes can be used to provide optimal performance. We have performed some computer simulations of the decoding algorithms we mentioned above and compared their performance to other existing ones. The results will be reported in the future.

7. References

[1] Bielecki, T., On integration with respect to fractional Brownian motions, Statistics and Probability Letters, to appear 1995

[2] Bielecki, T., On integration with respect to fractional Wiener sheet, preprint, 1995

[3] Bielecki, T., Chen, J. and Yau, Stephen S. T., Random wavelet transformation and its properties, in Wavelet Applications in Signal and Image Processing II, eds. A. F. Laine and M. A. Unser, SPIE Proc. vol. 2303 (1994)

[4] Chui, C., **An Introduction to Wavelets**, Academic Press, 1992

[5] Daubechies, I., Orthonormal basis of compactly supported wavelets, Commun. Pure Appl. Math., vol. 41, pp. 909-996 (1988)

[6] Ehrhard, D., Achieving the designed error capacity in decoding algebraic-geometric codes, IEEE Transactions on Information Theory, vol. 39, no. 3 (1993)

[7] Feng, G.-L. and Rao, T. R. N., Decoding algebraic-geometric codes up to the designed minimal distance, IEEE Transactions on Information Theory, vol. 39, no. 1 (1993)

[8] Gray, R. M., **Source coding theory**, Kluwer Academic Publishers, 1990

[9] Hartshorne, R. **Algebraic Geometry**, Springer-Verlag, 1990

[10] Haykin, S. **Digital Communications**, John Wiley & Sons, 1988

[11] Justesen, J., Larson, K. J., Jensen, H. E., and Hϕholdt, T. , Fast decoding of codes from algebraic plane curves, IEEE Transactions on Information Theory, vol. 38, no. 1 (1992)

[12] Keshner, M. S., $\frac{1}{f}$ Noise, Proceedings of the IEEE, vol. 70, no. 3, pp. 212-218 (1982)

[13] Kwong, M. K. and P. T. Peter Tang, W-matrices, nonorthogonal multiresolution analysis, and finite signals of arbitrary length, Argonne National Laboratory Preprint MCS-P449-0794.

[14] Kwong, M. K., MATLAB implementation of W-matrix multiresolution analyses, Argonne National Laboratory Preprint MCS-P462-0894.

[15] van Lint, **Introduction to Coding Theory**, Springer-Verlag, 1991

[16] van Lint and van der Geer, **Introduction to Coding and Algebraic Geometry**, Birkhauser, 1988

[17] Mallat, S., A theory for multiresolution signal decomposition: the wavelet representation, IEEE Transaction on Pattern Analysis and Machine Intelligence, vol. 11 pp. 674-693 (1989)

[18] Mandelbrot, B. B. and van Ness, T., Fractional Brownian motions, fractional noises and applications, SIAM Review, vol. 10, no. 4 (1968)

[19] Morgera, S. D. and Krishma, H., **Digital Signal Processing**, Academic Press, 1989

[20] Oppenheim, A. V. and Schafer, R. W., **Discrete-Time Signal Processing**, Prentice Hall, 1989

[21] SIAM News, vol. 27, number 2, Feb 1994.

[22] Song, L., Lectures on error correcting codes, algebraic geometric codes and the decoding algorithms, Lecture Notes, the University of Illinois at Chicago, fall 1993 and spring 1994

[23] Tsfaman, M. G. and Vlădut, S. G., **Algebraic Geometric Codes**, Kluwer Academic Publishers, 1991

[24] Verdó, S., and Wei, V. K., Explicit construction of optimal constant-weight codes for identification via channels, IEEE Transactions on Information Theory, vol. 39, no. 1 (1993)