WAVEFORM SIGNAL SHAPING USING WAVELET PARAMETERIZATIONS

Todd K. Moon
Electrical and Computer Engineering Department
Utah State University, Logan, UT 84322-4120

Krishna Kishor Noru
Sun Microsystems
901 San Antonio Rd, Palto Alto, CA 94303

ABSTRACT

We explore the idea of matching a scaling function - the basic building block of a wavelet function - to a desired spectrum. This would allow the scaling function to be used as the signal pulse for a digital communication system that is matched to the channel, avoiding problems such as energy loss or noise amplification due to spectral nulls. An unconstrained parameterization of the scaling function coefficients represents the scaling functions. This parameterization is adapted using gradient descent. Tests indicate that the adaptation is able to capture major features of a desired spectrum, including spectral nulls and major lobes.

KEY WORDS

Spectrum shaping; wavelets; adaptive spectrum

INTRODUCTION

In most digital communication systems, the signalling waveform shape is chosen primarily without regard to the channel through which the signal passes, except for the most pressing consideration of channel bandwidth. Thus, close attention is paid to spectral efficiency in signal selection, and waveforms such as GMSK or square-root raised cosine are frequently used [1]. However, there are other influences in a channel that may have bearing on the signal design. Consider the common circumstance in which the passband of the channel is not a perfect bandpass filter; more specifically, consider the case in which there is a significant notch in the passband. Ideally, less power should be expended at such frequencies, since they are attenuated by the channel. Or consider the case that the noise power spectrum density (PSD) is not white. Then the transmitted power should ideally be concentrated at those frequencies where the noise PSD is smallest. In fact, information theory prescribes exactly such a signal design in order to achieve capacity [2, Chapter 8]. Influences such as channel spectrum and noise spectrum are frequently dealt with at a receiver using equalizers or whitening filters, but these are suboptimal approaches. In conjunction with these methods there is the possibility of explicitly shaping the transmitted waveform to match some desired spectrum based on the channel and noise. In combination with other techniques, such as preceding to shape the correlation structure of the waveform, there is thus considerable flexibility available to the system designer in matching the signal to the channel. This moves us closer toward the view of digital communication expressed by Blahut [3, p. vii]:

We can give a whimsical definition of a digital communications system as a communication system designed to best use a given channel, and an analog communication system as one designed to best fit a given source. The spectrum of a well-designed digital communication waveform is a good match to the passband characteristics of the channel; the only way the source affects the spectrum of the waveform is by the bit rate.

In many channels, the response or noise PSD may change over time, so it is desirable to provide for adaptation of the transmitted waveform. This suggests the need for a parameterized signal that can be tuned, or adapted, to the changing channel. The goal of this paper is to introduce a parameterized family of waveforms which can be tuned to provide some degree of match to a desired spectrum $D(\omega)$. These waveforms could be used as the baseband waveforms for a digital communications system, where $D(\omega)$ represents the signal desired spectrum as determined by the channel response and the noise PSD. (The problem of determining $D(\omega)$ is not addressed by this paper; the literature of channel and spectral estimation introduces several approaches to the problem in a communications setting.) Another approach to this problem has also been explored in [4].

The family of waveforms explored here are instances of scaling functions (associated with wavelet functions). These offer many potentially interesting attributes in a communication setting, such as bandwidth efficiency, multi-scale transmission, and performance benefits in fast fading channels [5]. Using a parameterization due to Zou and Tewfik [6], scaling functions may be parameterized by only a few parameters. Hence, they make interesting candidates for baseband waveforms.

SUMMARY OF WAVELET PARAMETERIZATION

Scaling functions are the basic building blocks of dyadic wavelet transforms. A scaling function $\phi(t)$ is a low-pass function that satisfies the two-scale dilation equation [7, 6]

$$\phi(t) = \sum_{k=0}^{N-1} c_k \phi(2t-k). \tag{1}$$

Wavelets are built from scaling functions using a two-scale relation

$$\psi(t) = \sum_{k=0}^{N-1} d_k \phi(2t-k).$$

In order to satisfy orthogonality requirements of wavelet transforms, the sequences of coefficients $\{c_k\}$ must satisfy the the constraints

$$\begin{aligned} \sum_{k=0}^{N-1} c_k &= 2 \\ \sum_{k=0}^{N-1} c_k c_{k+2m} &= 2\delta_m \\ \sum_{k=)}^{N-1} (-1)^k k^m c_k &= 0 \end{aligned} \tag{2}$$

for $m = 1,...,M - 1$, where M is said to be the number of vanishing moments of the scaling function. The number of coefficients N must be even. Also, the $\{d_k\}$ coefficients satisfy

$$d_k = (-1)^k c_{N-k}; \tag{3}$$

that is, they are reverse in order, and alternate in sign. Subject to (2) and (3), the set of functions $\{2^{j/2}\psi(2^j t - k); j, k \in \mathbb{Z}\}$ forms an orthonormal basis for $L^2(\mathbb{R})$ [7] . As the coefficients $\{c_k\}$ are changed, subject to (2), various scaling and wavelet functions are produced.

Suppose now that some desired spectrum $D(\omega)$ is specified, and it is desired to determine a scaling function $\phi(t)$ whose transform $\hat{\phi}(\omega)$ matches the desired spectrum as closely as possible, so that $||D(\omega) - \hat{\phi}(\omega)||$ is minimized. We will use the L_2 norm, so that

$$||D(\omega) - \hat{\phi}(\omega)|| = \int_0^\infty |D(\omega) - \hat{\phi}(\omega)|^2 \, d\omega$$

(where only the positive frequency terms are used because of an assumed symmetry). Let c denote the vector of N parameters, c = $[c_0, c_1,..., c_{N-1}]^T$. We will represent the parameterization of the scaling function by $\hat{\phi}(\omega; \mathbf{c})$. The minimization problem can be stated as

$$\min_{\mathbf{c}} \int_0^\infty |D(\omega) - \hat{\phi}(\omega; \mathbf{c})|^2 \, d\omega,$$

subject to the conditions in (2).

Satisfying the constraints in as stated in (2) is quite complicated. However, there is a different parameterization of c which automatically enforces the constraints. Let

$$H(z) = \sum_{k=0}^{N-1} c_k z^{-k}$$

and

$$G(z) = \sum_{k=0}^{N-1} d_k z^{-k}$$

be the discrete-time filters corresponding to the coefficient sequences. The polyphase representations of $H(z)$ and $G(z)$ are obtained by collecting the even- and odd-exponent terms together as

$$H(z) = E_{00}(z^2) + z^{-1} E_{01}(z^2)$$

$$G(z) = E_{10}(z^2) + z^{-1} E_{11}(z^2).$$

This can be written in matrix form as

$$\begin{bmatrix} H(z) \\ G(z) \end{bmatrix} = \begin{bmatrix} E_{00}(z^2) & E_{01}(z^2) \\ E_{10}(z^2) & E_{11}(z^2) \end{bmatrix} \begin{bmatrix} 1 \\ z^{-1} \end{bmatrix} \triangleq \mathbf{E}(z^2) \begin{bmatrix} 1 \\ z^{-1} \end{bmatrix}. \quad (4)$$

It can be argued [6, 8, 9] that if the matrix $\mathbf{E}(z)$ is lossless, then constraints (2) are satisfied, where a lossless matrix $\mathbf{E}(z)$ is a stable matrix for which

$$\tilde{\mathbf{E}}(z)\mathbf{E}(z) = cI$$

and where

$$\tilde{\mathbf{E}}(z) = \mathbf{E}_*^T(z^{-1}),$$

and where $\mathbf{E}_*$ indicates conjugation of the coefficients without conjugating z. If we restrict our attention to purely real filters, we have

$$\tilde{\mathbf{E}}(z) = \mathbf{E}^T(z^{-1}).$$

A parameterization for a lossless matrix is thus a parameterization for a scaling function. It can also be shown [10, 11] that every 2 x 2 lossless matrix $\mathbf{E}$(z) of degree k can be written as

$$\mathbf{E}(z) = V_k(z)V_{k-1}(z)\cdots V_1(z)V_0$$

where

$$V_i(z) = I - \mathbf{v}_i\mathbf{v}_i^H + z^{-1}\mathbf{v}_i\mathbf{v}_i^H,$$

where ^{H} indicates conjugate transpose, and where $\mathbf{v}_i$ is a unit vector. The matrix V_0 is a constant unitary matrix, which we will take as

$$V_0 = \begin{bmatrix} \cos\theta_0 & -\sin\theta_0 \\ \sin\theta_0 & \cos\theta_0 \end{bmatrix}$$

Without loss of generality, we can take

$$\mathbf{v}_i = \begin{bmatrix} \cos\theta_i \\ \sin\theta_i \end{bmatrix}.$$

Since $H(z)$ has N coefficients, the polynomials in $\mathbf{E}(z)$ of (4) has $N/2$ coefficients (degree $N/2$ - 1), so it can be parameterized as

$$\mathbf{E}(z) = V_{N/2-1}(z)V_{N/2-2}(z)\cdots V_1(z)V_0$$

Using (4), we now have a parameterization of the coefficients of $H(z)$ in terms of $\theta = \{\theta_0, \theta_1, \ldots, \theta_{N/2-1}\}$, and these parameters are *unconstrained*. The vanishing moment condition can be shown [6] to lead to the condition that $\theta_0 = 3\pi/4$. There are thus $N/2$ - 1 free parameters $\theta_0, \theta_1, \ldots, \theta_{N/2-1}$ to describe the scaling function with N coefficients. We will let the scaling function θ parameterized by the coefficient set θ be denoted by $\theta(t; \theta)$, and its Fourier transform by $\hat{\phi}(\omega; \boldsymbol{\theta})$.

SPECTRAL PROPERTIES OF SCALING FUNCTIONS

Let $m_0(z) = \sum_{k=0}^{N-1} c_k z^k$ denote the *Z*-transform of the scaling coefficient sequence. By (1) and the properties of Fourier transforms, if follows that

$$\hat{\phi}(\omega) = m_0(e^{-j\omega/2})\hat{\phi}(\omega/2)$$

and, by induction,

$$\hat{\phi}(\omega) = \prod_{l=1}^{n} m_0(e^{-j\omega/2^l})\hat{\phi}(\omega/2^n) \overset{n\to\infty}{\to} \prod_{l=1}^{\infty} m_0(e^{-j\omega/2^l})$$

since $\hat{\phi}(0) = \int \phi(t)\,dt = 1$. In practice, the infinite product is approximated by choosing some finite upper limit.

For the problem of matching a scaling function to a desired spectrum $D(\omega)$, we want to minimize

$$J(\boldsymbol{\theta}) = \int_0^\infty |D(\omega) - \hat{\phi}(\omega;\boldsymbol{\theta})|^2\,d\omega. \tag{5}$$

GRADIENT DESCENT OPTIMIZATION

Let $\boldsymbol{\theta}^{[p]} = \{\theta_1^{[p]}, \theta_2^{[p]}, \ldots, \theta_{N/2-1}^{[p]}\}$ denote the parameter set at the *p*th iteration of the algorithm. The gradient descent update of θ is obtained from

$$\boldsymbol{\theta}^{[p+1]} = \boldsymbol{\theta}^{[p]} - \mu\nabla_{\boldsymbol{\theta}} J(\boldsymbol{\theta}^{[p]})$$

where μ is a step size parameter. The gradient can be computed from The gradient can be computed as follows:

$$\begin{aligned}
\nabla_\theta J(\boldsymbol{\theta}) &= \int_0^\infty \nabla_\theta \left(|D(\omega)|^2 - 2\,\mathrm{Re}(D(\omega)\hat{\phi}(\omega;\boldsymbol{\theta})) + |\hat{\phi}(\omega;\boldsymbol{\theta})|^2\right)\,d\omega \\
&= -2\int_0^\infty \nabla_\theta\,\mathrm{Re}(D(\omega)\hat{\phi}(\omega;\boldsymbol{\theta}))\,d\omega & (6)\\
&= -2\int_0^\infty \mathrm{Re}(D(\omega)\nabla_\theta\hat{\phi}(\omega,\boldsymbol{\theta}))\,d\omega & (7)
\end{aligned}$$

where (6) follows since the energy of ϕ does not depend on the coefficients and (7) follows since $D(\omega)$ is not a function of θ. Then the gradient can be computed as,

$$\nabla_{\boldsymbol{\theta}}\hat{\phi}(\omega;\boldsymbol{\theta}) = \nabla_{\boldsymbol{\theta}} \prod_{l=1}^{\infty} m_0(e^{-j\omega/2^l};\boldsymbol{\theta}), \tag{8}$$

To be practically implemented, this product must be truncated at a finite number of terms, *P*. Let

$$\hat{\phi}_P(\omega;\boldsymbol{\theta}) = \prod_{l=1}^{P} m_0(e^{-j\omega/2^l};\boldsymbol{\theta}). \tag{9}$$

Then

$$\nabla_{\boldsymbol{\theta}} \hat{\phi}_P(\omega; \boldsymbol{\theta}) = \nabla_{\boldsymbol{\theta}} \prod_{l=1}^{P} m_0(e^{-j\omega/2^l}; \boldsymbol{\theta}).$$

Because of the product, it is convenient to using the logarithmic differentiation rule $\nabla f = f \nabla \log f$ to obtain

$$\nabla_{\theta} \hat{\phi}_P(\omega; \boldsymbol{\theta}) = \hat{\phi}_P(\omega; \boldsymbol{\theta}) \sum_{l=1}^{P} \frac{\nabla_{\theta} m_0(e^{-j\omega/2^l}; \boldsymbol{\theta})}{m_0(e^{-j\omega/2^l}; \boldsymbol{\theta})} \tag{10}$$

$$= \hat{\phi}_P(\omega; \boldsymbol{\theta}) \sum_{l=1}^{P} \frac{\mathcal{P}(e^{-j\omega/2^l}; \boldsymbol{\theta})}{m_0(e^{-j\omega/2^l}; \boldsymbol{\theta})} \tag{11}$$

where

$$\mathcal{P}(z) = [p_0(z), p_1(z), \ldots, p_{N/2-1}(z)]^T$$

with

$$p_l(z) = \sum_k \frac{\partial c_k(\boldsymbol{\theta})}{\partial \theta_l} z^k$$

The integral in (7) may be discretized for rapid computation, and written as

$$\nabla_{\boldsymbol{\theta}} J(\boldsymbol{\theta}) \approx -2 \sum_{d=1}^{N_p} \mathrm{Re}(D(\omega_d) \nabla_{\theta} \hat{\phi}(\omega_d, \boldsymbol{\theta}))(\omega_d - \omega_{d-1}) \tag{12}$$

where N_p represents the number of discretization intervals and ω_d represents the frequency points of interest. By iteratively running the loop either till $J(\theta)$ has a satisfactory minimal value or until a minimum is reached, a good estimate of the parameter θ can be made.

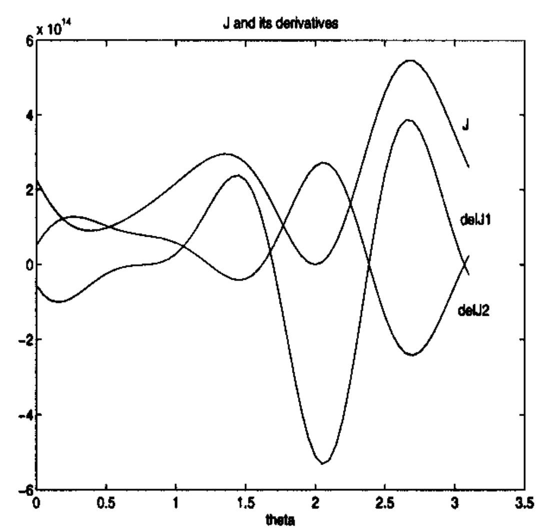


Figure 1: A plot of $J(\theta)$ and its derivatives.

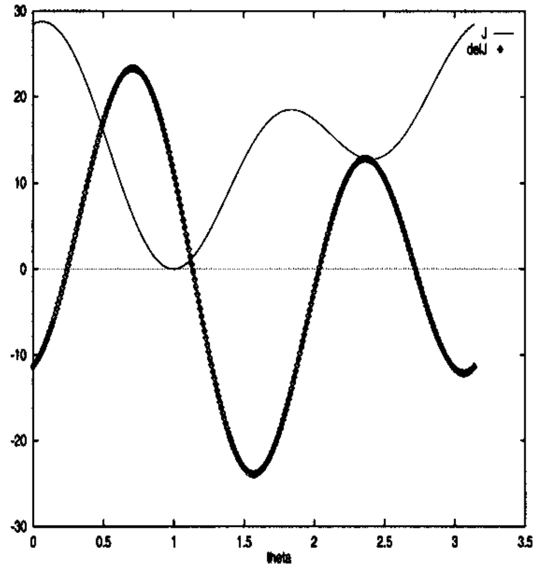


Figure 2: A plot of $J(\theta)$ and its derivative. The plot J represents $J(\theta)$ and delJ1 the derivative of $J(\theta)$ with respect to the parameter θ_1. The plot of J shows more than one minimum.

TESTING

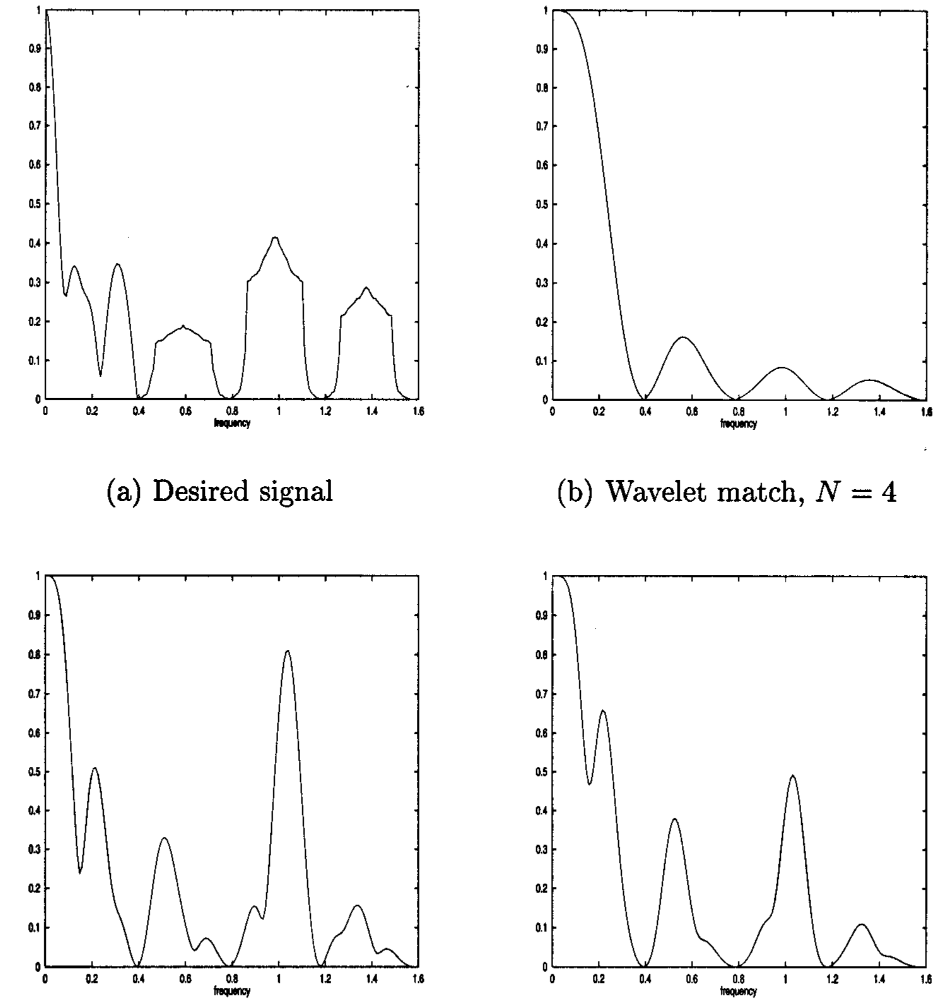
Because of truncation in the infinite product, the derivative is only an approximation to the true derivative. This is demonstrated in figure 1, which shows $J(\theta)$ and its derivatives with respect to θ_1 and θ_2. In this figure, the desired spectrum is obtained from a four-coefficient scaling function. The horizontal axis in the figure is the θ_2 parameter, and the other parameter is fixed. When $\theta_2 = 2$, $J(\theta) = \theta$, but the computed derivatives, indicated by delJ1 and delJ2, are not zero at the point of minimum, due to truncation.

As with most gradient descent methods, there is always the potential for the algorithm to converge to a local minimum. This combines with the approximate gradient problem to lead to only approximate solutions. As an example, figure 2 shows $J(\theta)$ (again where $D(\omega)$ is a scaling function) as a function of a parameter θ and also the derivative with respect to θ. As observed before, the gradient is not zero at the points of minimum (because of the approximation). If the starting value of $\theta = 0.9$ is used, the gradient descent algorithm converges to a final value of $\theta = 1.135$ (the value where the derivative is zero), and if the starting point of $\theta = 2.2$ is chosen, the gradient descent algorithm converges to $\theta = 2.718$.

Figure 4(a) illustrates a desired spectrum $D(\omega)$ generated at random with only one side lobe. With $N = 6$ parameters, starting from an initial θ of (3.1, 2.0), the gradient descent algorithm gave a final solution of $\theta = (2.0, 0.246)$, producing the scaling function shown in figure 4(b). After convergence, $J(\theta) = 13.14$. Note that the algorithm has found a function which captures the major features of the desired waveform, including the point of zero spectrum and the presence of the sidelobe, but has not captured all the fine detail.

Figure 3(a) illustrates another randomly selected desired spectrum $D(\omega)$ having multiple lobes. Figure 3(b) illustrates the scaling function spectrum found by gradient descent when $N = 4$, starting from $\theta = 3.1$ and converging to $\theta = 0.322$. Figure 3(c) illustrates the spectrum when $N = 6$, starting from θ (3.1, 1.0) and converging to 2.969, 0.621); the fitness function was $J(\theta) = 11.60$. Figure 3(d) illustrates the spectrum when $N = 8$ starting from $\theta = (3.1, 1.0, 2.5)$ and converging to (0. 289, 0.445, 2.62), with $J(\theta) = 15.0186$.

In each case, the major features of the desired spectrum are found, including the locations of the zero response frequencies and approximations to the peaks.



(a) Desired signal

(b) Wavelet match, $N = 4$

(c) Wavelet match, $N = 6$

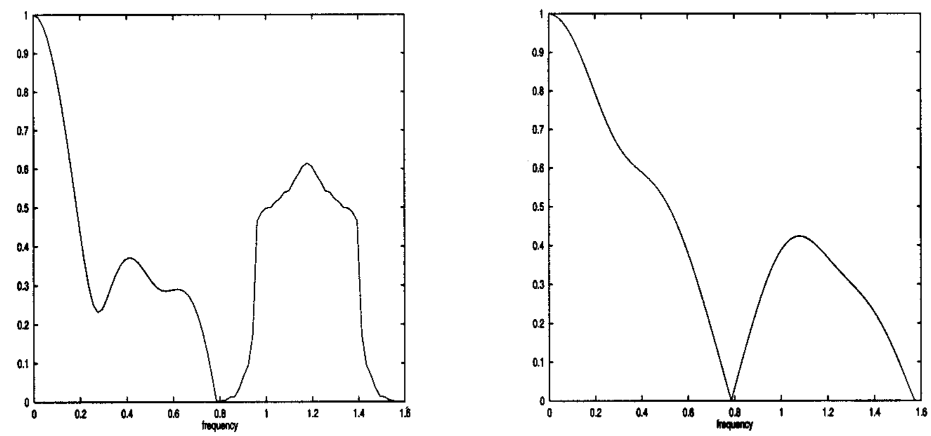
(d) Wavelet match, $N = 8$

Figure 3: A desired signal spectrum generated randomly for checking the performance of gradient descent algorithm, and matched spectra for values of N

SUMMARY

A method of matching scaling function waveforms to desired spectra has been discussed using gradient descent on an unconstrained parameterization of scaling functions. Tests indicate that due to approximations in the spectrum the gradient descent produces results that are biased and may converge only to a local minimum. Nevertheless, tests with randomly generated desired spectra indicate that the major features of a spectrum are obtained when the algorithm is used.

Application of this may be appropriate to slowly varying channels with spectral nulls, for which transmission of energy at spectral null frequencies would be wasteful and for which equalization would be difficult. Because it is a gradient descent approach, the method may be appropriate for an on-line adaptation technique when used in conjunction with an on-line channel estimator.



(a) Desired (b) Obtained

Figure 4: A desired spectrum generated randomly with one sidelobe, and its wavelet match using gradient descent. side lobe

References

[1] J. C. Proakis and M. Salehi, *Communication Systems Engineering*. Prentice-Hall, 1994.

[2] R. G. Gallager, *Information Theory and Reliable Communication*. New York: Wiley, 1968.

[3] R. E. Blahut, *Digital Transmission of Information*. Reading, MA: Addison Wesley, 1990.

[4] A. H. Tewfik, D. Sinha, and P. Jorgensen, "On the Optimal Choice of a Wavelet for Signal Representation," *IEEE Trans. Information Theory*, vol. 38, pp. 747-765, March 1992.

[5] C. Lo and T. Moon, "Multiscale wavelet signalling and its performance in a ricean fast fading channel (submitted for publication)," *J. Comm and Networks*, 1999.

[6] H. Zou and A. H. Tewfik, "Parametrization of Compactly Supported Orthonormal Wavelets," *IEEE Trans. Signal Processing*, vol. 41, pp. 1428-1431, March 1993.

[7] I. Daubechies, *Ten Lectures on Wavelets*. Philadelphia, PA: SIAM, 1992.

[8] I. Daubechies, "Orthonormal Bases of Compactly Supported Wavelets," *Com. on Pure and Appl. Math.*, vol. XLI, pp. 909-996, 1988.

[9] P. P. Vaidyanathan, "Multirate Digital Filters, Filter Banks, Polyphase Networks, and Applications: A Tutorial," *Proc. IEEE*, vol. 78, pp. 56-93, Jan. 1990.

[10] P. Vaidyanathan and Z. Doğanata, "The role of lossless systems in modern digital signal processing: a tutorial," *IEEE Trans. Education*, Aug. 1989.

[11] P. Vaidyanathan, T. Nguyen, Z. Doğanata, and T. Sarmaki, "Improved search for design of perfect reconstruction FIR QMF banks with lossless polyphase structures," *IEEE Trans. Acoust., Speech, Signal Processing*, July 1989.