# WAVELET AND FOOTPRINT SAMPLING OF SIGNALS WITH A FINITE RATE OF INNOVATION


*Pier Luigi Dragotti*

Electrical and Electronic Engineering
Imperial College London
Exhibition Road, London SW7-2AZ, UK
p.dragotti@imperial.ac.uk

*Martin Vetterli*

Audio-Visual Communications Laboratory
École Polytechnique Fédérale de Lausanne
CH - 1015 Lausanne, Switzerland.
martin.vetterli@epfl.ch



### ABSTRACT

In this paper, we consider classes of not bandlimited signals, namely, streams of Diracs and piecewise polynomial signals, and show that these signals can be sampled and perfectly reconstructed using wavelets as sampling kernel. Due to the multiresolution structure of the wavelet transform, these new sampling theorems naturally lead to the development of a new resolution enhancement algorithm based on wavelet footprints [2]. Preliminary results show the potentiality of this algorithm.


## 1. INTRODUCTION

A critical element in modern signal processing and communication is sampling. Most continuous-time phenomena are analyzed through sampling. Often, the original continuous-time signal $x(t)$ is filtered before sampling and this filtering may be due to the acquisition device or may be a design choice. Let $h(t)$ be the impulse response of this filter. Then, the uniform sampling of $x(t)$ with sampling interval $T$ leads to samples $y_n$ given by

$$y_n = \langle h(t-nT), x(t)\rangle = \int_{-\infty}^{\infty} h(t-nT)x(t)dt.$$

The key problem is to find the best way to reconstruct $x(t)$ from its samples. If $x(t)$ is bandlimited, than the Shannon sampling theorem states the conditions to reconstruct $x(t)$ from $y_n$'s.

Recently, it was shown that it is possible to develop sampling schemes for classes of signals that are not band-limited [8]. In particular, it was shown that it is possible to sample streams of Diracs and piecewise polynomial signals using a sinc or a Gaussian kernel. The common feature of these signals is that they have a parametric representation with a finite number of degrees of freedom. This number of degrees of freedom is called rate of innovation. Thus, streams of Diracs and piecewise polynomial signals are signals with a finite rate of innovation.

In this paper, we extend the results of [8] and show that streams of Diracs and piecewise polynomial signals can be sampled and perfectly reconstructed using wavelets as sampling kernel. Due to the multiresolution structure of wavelets, these new results naturally lead to a new algorithm for resolution enhancement. This algorithm is based on the notion of wavelet footprints which was introduced in [2]. For an excellent review on sampling we refer to [5]. Some pioneering works on sampling with the wavelet transform can be found in [6].

The paper is organized as follows: The next section presents a brief review of the wavelet transform. In Section 3, we present new sampling theorems for signal with a finite rate of innovation. In Section 4 we provide an interpretation of these sampling results in terms of resolution enhancement and we present a footprint-based algorithm for resolution enhancement. We conclude in Section 5.

## 2. THE WAVELET TRANSFORM

This section presents a brief review of the wavelet transform. For a more detailed treatment, we refer the reader to [1, 7, 4, 3].

Consider a wavelet function $\psi(t)$ that generates a basis of $L_2(\mathbb{R})$. That is, assume that $\psi(t)$ satisfies the admissibility condition and that the set of its dilated and shifted versions $\psi_{m,n}(t) = \frac{1}{2^{m/2}}\psi(2^{-m}t-n) \quad m,n \in \mathbb{Z}$ forms a basis of $L_2(\mathbb{R})$. The discrete wavelet transform is a unique and stable decomposition of any finite energy signal $x(t)$ in terms of $\{\psi_{m,n}\}_{m\in\mathbb{Z},n\in\mathbb{Z}}$ or

$$x(t) = \sum_{m=-\infty}^{\infty}\sum_{n=-\infty}^{\infty} d_{m,n}\psi_{m,n}(t). \tag{1}$$

The wavelet coefficients $d_{m,n}$ are given by $d_{m,n} = \langle x(t), \tilde{\psi}_{m,n}(t)\rangle$ with $\tilde{\psi}_{m,n}(t)$ such that $\langle \tilde{\psi}_{m,n}(t), \psi_{j,k}(t)\rangle = \delta_{m-j}\cdot\delta_{n-k}$. To be more precise, the set $\{\tilde{\psi}_{m,n}\}_{m\in\mathbb{Z},n\in\mathbb{Z}}$ represents the dual basis of $\{\psi_{m,n}\}_{m\in\mathbb{Z},n\in\mathbb{Z}}$. If $\{\psi_{m,n}\}_{m\in\mathbb{Z},n\in\mathbb{Z}}$ is an orthonormal basis of $L_2(\mathbb{R})$, then the two sets $\{\psi_{m,n}\}_{m\in\mathbb{Z},n\in\mathbb{Z}}$ and $\{\tilde{\psi}_{m,n}\}_{m\in\mathbb{Z},n\in\mathbb{Z}}$ coincide.

The double sum in (1) clearly shows the multiresolution structure of the wavelet transform. Since the wavelet function $\psi(t)$ has zero average, each term $d_{m,n}$ measure a local variation of $x(t)$ at resolution $2^m$ and the partial sum

$$x_{J+1}(t) = \sum_{m=J+1}^{\infty}\sum_{n=-\infty}^{\infty} d_{m,n}\psi_{m,n}(t) \tag{2}$$

represents an approximation of $x(t)$ at resolution $2^{J+1}$. The completeness of $\{\psi_{m,n}\}_{m\in\mathbb{Z},n\in\mathbb{Z}}$ ensures that by adding details to $x_{J+1}(t)$ at finer and finer resolutions we eventually recover $x(t)$.

The approximation function $x_{J+1}(t)$ can be expressed in terms of shifted versions of a different function $\varphi(t)$ called the scaling function. That is

$$x_{J+1}(t) = \sum_{n=-\infty}^{\infty} y_{J,n}\varphi_{J,n}(t) \tag{3}$$

0-7803-8484-9/04/$20.00 ©2004 IEEE

with the usual assuption that $\varphi_{J,n}(t) = 1/2^{J/2}\varphi(t/2^J - n)$. Thus, by combining (1), (2) and (3), we see that it is possible to represent any function in $L_2(\mathbb{R})$ as a combination of wavelets and scaling functions or

$$x(t) = \sum_{n=-\infty}^{\infty} y_{J,n}\varphi_{J,n}(t) + \sum_{m=-\infty}^{J}\sum_{n=-\infty}^{\infty} d_{m,n}\psi_{m,n}(t). \quad (4)$$

The scaling coefficients $y_{J,n}$ tend to measure the local regularity of $x(t)$ at scale $2^J$. Therefore, the term $\sum_{n=-\infty}^{\infty} y_{J,n}\varphi_{J,n}(t)$ represents a coarse version of $x(t)$ as opposed to the detail version provided by the wavelets in the last term of (4).

The wavelet function and the scaling function are intimately related and their link does not reduce to the expansion showed in (4). Indeed, the scaling function represents the basic element in the construction of a wavelet basis and many properties of wavelets can be inferred directly from the scaling function. In particular, an important and well known property of the wavelet transform is that of the vanishing moments. We say that a wavelet has $K$ vanishing moments if

$$\int_{-\infty}^{\infty} t^k \tilde{\psi}(t)dt = 0, \qquad k = 0, 1, ..., K-1.$$

This vanishing moments property translates directly into the polynomial approximation property of the scaling function. More precisely, a wavelet has $K$ vanishing moments if and only if its corresponding scaling function can reproduce polynomials of maximum degree $K-1$, that is,

$$\sum_{n\in\mathbb{Z}} c_{k,n}\varphi(t-n) = t^k \qquad k = 0, 1, ..., K-1. \quad (5)$$

In the next section, we will use the properties of the scaling function to present new sampling results for classes of signal with a finite rate of innovation. In addition, we will use the link between wavelets and scaling functions and the multiresolution nature of wavelets to give an interpretation of these sampling results in terms of resolution enhancement.

## 3. WAVELET SAMPLING OF SIGNALS WITH FINITE RATE OF INNOVATION

In this section, we consider scaling functions of compact support $L$, that is, $\varphi(t) \neq 0$ for $t \in [-L/2, L/2]$ where $L$ is for simplicity an integer; and we assume that a linear combination of $\varphi(t)$ can reproduce polynomials of degree $K-1$. We concentrate on one class of signals, namely streams of Diracs. In particular, we show that the sampling problem reduces to the problem of solving a system of polynomial equations and that there is a trade-off between the complexity of this set of equations and the local rate of innovation of the sampled signal. Eventhough we focus only on streams of Diracs, most of the results which are valid for this class of signals can be extended to piecewise polynomial signals.

Consider a stream of Diracs $x(t) = \sum_{n\in\mathbb{Z}} a_n\delta(t-t_n)$ and $t \in \mathbb{R}$ and assume that there is at most one Dirac in an interval of length $LT$. It follows

**Proposition 1** *Given is a scaling function $\varphi(t)$ of compact support $L$ and that can reproduce polynomials of maximum degree one. An infinite-length stream of Diracs $x(t) = \sum_{n\in\mathbb{Z}} a_n\delta(t-t_n)$ is uniquely determined from the samples defined by $y_n = \langle\varphi(t/T-n), x(t)\rangle$ if and only if there is at most one Dirac in an interval of length $LT$.*

*Proof:* We first show how to localize a Dirac in an interval of size $T$, then we show how to find the exact location and amplitude of that Dirac.

Let $T = 1$ and let the support of $\varphi(t)$ be $L$, assume the signal is known for $t \leq n - L/2$. If there is no Dirac in $[n-L/2, n+L/2]$ then $y_n = 0$. If there is one Dirac in that interval (call it $a_k\delta(t-t_k)$), then $y_n \neq 0$. Now, consider the inner product $y_{n-L+1}$ if there is no Dirac in the interval $[n-3L/2+1, n-L/2]$ and $y_{n-1} \neq 0$, then the dirac $a_k\delta(t-t_k)$ is in the interval $[n-L/2, n-L/2+1]$. If a Dirac was already found in the interval $[n-3L/2+1, -L/2n]$ (recall that $x(t)$ is known for $t \leq n-L/2$) or if $y_{n-1} = 0$ then $a_k\delta(t-t_k)$ cannot be in $[n-L/2, n-L/2+1]$, but must be in $[n-L/2+1, n+L/2]$. We then need $y_{n-L+2}$ to see if $a_k\delta(t-t_k)$ is in $[n-L/2+1, n-L/2+2]$. The process is iterated until we find an interval of size $T$ where we know $a_k\delta(t-t_k)$ is. Assume $y_n$ $n = 0, 1, ..., L-1$ are the inner products that overlap this interval. Since the scaling function has compact support $L$ and there is at most one Dirac in an interval of length $L$, we are sure that only $L$ inner products overlap $a_k\delta(t-t_k)$ and no other Diracs are in the same inner products. Therefore using partition of unity and equation (5), we have that

$$a_k = \sum_{n=0}^{L-1} y_n \quad (6)$$

and

$$t_k = (\sum_n c_{1,n}y_n)/a_k \quad (7)$$

where the coefficients $c_{1,n}$ are known and given by (5).

□

In equations (6) and (7), we have used the fact that, in the vicinity of $t_k$, the scaling function is reproducing polynomial of degree zero and one respectively. In fact, we have that

$$\begin{aligned} \textstyle\sum_{n=0}^{L-1} y_n &= \langle a_k\delta(t-t_k), \textstyle\sum_{n=0}^{L-1}\varphi(t-n)\rangle \\ &= \textstyle\int_{-\infty}^{\infty} a_k\delta(t-t_k)(\sum_{n=0}^{L-1}\varphi(t-n))dt \quad (8) \\ &= a_k\textstyle\sum_{n=0}^{L-1}\varphi(t_k-n) = a_k \end{aligned}$$

where in the last equality we have used the property that the sum of the translated versions of $\varphi(t)$ is constant and equal to 1 in $t_k$. Likewise, we have that

$$\sum_{n=0}^{L-1} c_{1,n}y_n = a_k\sum_{n=0}^{L-1} c_{1,n}\varphi(t_k-n) = a_k t_k \quad (9)$$

where in the last equality we used the polynomial approximation property (5). Figure 1 illustrates this result with a simple example.

Finally, it is worth pointing out that the scaling functions that generate some of the most commonly used wavelets such as Daubechies wavelets and Splines, satisfy the hypotheses of the theorem and can, therefore, be used to sample streams of Diracs.

The proposition above has shown conditions under which we can sample streams of Diracs. The reconstruction algorithm relies on the ability of the scaling functions to reproduce polynomials of degree one. However, we need to assume that there is at most one Dirac in an interval of size $LT$. We can loosen this condition by assuming that $\varphi(t)$ can reproduce higher order polynomials. In particular we have that:

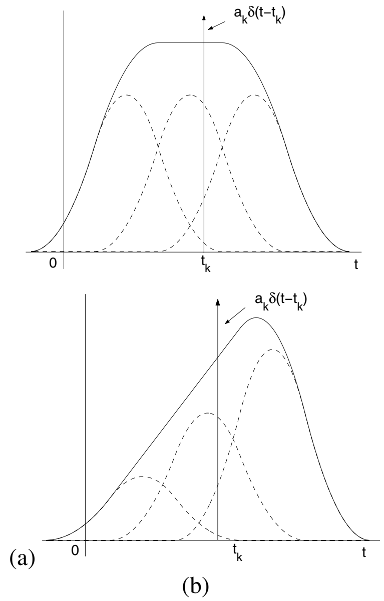


**Fig. 1**. Illustration of the sampling result of Proposition 1 using B-splines of degree two. In this case $L$ is equal to three and only three translated versions of the scaling function overlap the Dirac. The three dashed functions in part (a) and (b) are the three B-splines overlapping the Dirac. In part (b), they are opportunely weighted to reproduce a degree-one polynomial. The two solid-line functions in (a) and (b) represents $\varphi(t)+\varphi(t-1)+\varphi(t-2)$ and $c_{1,0}\varphi(t)+c_{1,1}\varphi(t-1)+c_{1,2}\varphi(t-2)$ respectively. Because of the polynomial reproduction property of the scaling function, the following is true: $y_0+y_1+y_2=\int_{-\infty}^{\infty}a_k\delta(t-t_k)(\varphi(t)+\varphi(t-1)+\varphi(t-2))dt=a_k$ where in the last equality, we have used the fact that around $t_k$ the sum of the scaling functions is constant and equal to one. Similarly, $c_{1,0}y_0+c_{1,1}y_1+c_{1,2}y_2=a_kt_k$ as illustrated in Figure 1(b).

**Proposition 2** *Given is a scaling function $\varphi(t)$ that can reproduce polynomials of maximum degree three and of compact support $L$. An infinite-length stream of Diracs $x(t)=\sum_{n\in\mathbb{Z}}a_n\delta(t-t_n)$ is uniquely determined from the samples defined by $y_n=\langle\varphi(t/T-n),x(t)\rangle$, if and only if there are at most two Diracs in an interval of length $2LT$.*

*Proof:* In a way similar to the one presented in Proposition 1, we can find the interval that contains the two Diracs (call them $a_0\delta(t-t_0)+a_1\delta(t-t_1)$ with the assumption that $t_0\leq t_1$). Assume $y_n$ $n=0,1,...,L'$ are the only inner products that overlap the two Diracs. The following is true

$$\sum_n y_n=a_0+a_1 \tag{10}$$

$$\sum_n c_{1,n}y_n=a_0t_0+a_1t_1 \tag{11}$$

$$\sum_n c_{2,n}y_n=a_0t_0^2+a_1t_1^2 \tag{12}$$

$$\sum_n c_{3,n}y_n=a_0t_0^3+a_1t_1^3 \tag{13}$$

This is a system of four polynomial equations in four unknowns $(a_0,a_1,t_0,t_1)$, we need to show that it admits only one solution.

It is easy to see that after few manipulations this system can be written in triangular form as follows

$$s_0t_1^2-s_0kt_1+s_1k-s_2=0, \tag{14}$$

$$t_0=k-t_1, \tag{15}$$

$$a_1=(s_1-s_0t_0)/(t_1-t_0), \tag{16}$$

$$a_0=s_0-a_1 \tag{17}$$

with $s_0=\sum_n y_n$, $s_1=\sum_n c_{1,n}y_n$, $s_2=\sum_n c_{2,n}y_n$, $s_3=\sum_n c_{3,n}y_n$ and $k=(s_1s_2-s_0s_3)/(s_1-s_0s_2)$. Thus, we can solve equation (14) in $t_0$ and then substitute the values of $t_0$ in the other equations to find the exact values of $t_1$, $a_1$, $a_0$. Equation (14) has two solutions, therefore the whole system has apparently two possible sets of solutions. However, notice that the role of $t_0$ and $t_1$ in equations (14) and (15) can be exchanged. This means that, the two pairs of solutions that we obtain for $t_0$ and $t_1$ are symmetric. That is, if $t_0$ admits solutions $\alpha$ and $\beta$, than the corresponding solutions for $t_1$ are $\beta$ and $\alpha$ respectively. Therefore, following our convention that $t_0\leq t_1$ and assuming $\alpha\leq\beta$, we have that $t_0=\alpha$ and $t_1=\beta$ and the complete system admits only one solution. □

Finally, it is also possible to show that [1]

**Proposition 3** *An infinite-length stream of fixed amplitude Diracs $x(t)=\sum_{n\in\mathbb{Z}}\delta(t-t_n)$ is uniquely determined from the samples defined by $y_n=\langle\varphi(t/T-n),x(t)\rangle$, where $\varphi(t)$ is a scaling function of compact support L and that can reproduce polynomials of maximum degree $K$, if and only if there are at most $K$ Diracs in an interval of length $KLT$.*

Before concluding this section, we would like to highlight how to extend these sampling results to the case of piecewise polynomial signals. By differentiation, a piecewise polynomial signal can be reduced to a stream of Diracs. Thus, using integration by parts, one can sample piecewise polynomial signals using derivative of the scaling functions. We omit this proof for lack of space.

## 4. SAMPLING AND RESOLUTION ENHANCEMENT WITH FOOTPRINTS

In this section we investigate the use of footprints to reconstruct or to increase the resolution of a sampled signal. Wavelet footprints were introduced in [2].

We have seen that a signal $x(t)\in L_2(\mathbb{R})$ can be decomposed in terms of wavelets and scaling functions or

$$x(t)=\sum_{n=-\infty}^{\infty}y_{J,n}\varphi_{J,n}(t)+\sum_{m=-\infty}^{J}\sum_{n=-\infty}^{\infty}d_{m,n}\psi_{m,n}(t). \tag{18}$$

Now, assume that $x(t)$ and $\varphi(t)$ satisfies the hypotheses of the theorems in Section 3. That is, $x(t)$ is a stream of Diracs or a piecewise polynomial signal with a finite rate of innovation, and $\varphi(t)$ is a compact support scaling function that can reproduce polynomials of a certain degree. Then the sampling theorems of the previous section ensure that, for a proper choice of $J$, the inner products $y_{J,n}$ of equation (18) are sufficient to characterize $x(t)$ or, in other words, that the finite resolution version $x_{J+1}(t)=$

[1]We omit this proof due to the lack of space.

$\sum_{n=-\infty}^{\infty} y_{J,n}\varphi_{J,n}(t)$ is sufficient to reconstruct the signal exactly. This means that by knowing $x(t)$ at a finite resolution, we can infer the value of the wavelet coefficients $d_{m,n}$ with $m \leq J$ and, therefore, arbitrary increase the resolution of our approximation to eventually recover the original signal.

These sampling results can also by interpreted in terms of footprints. Consider, for instance, the case where $x(t)$ is a stream of Diracs, that is, $x(t) = \sum_{k\in\mathbb{Z}} a_k\delta(t - t_k)$. We know that we can write $x(t)$ as

$$x(t) = \sum_{n=-\infty}^{\infty} y_{J,n}\varphi_{J,n}(t) + \sum_{k\in\mathbb{Z}} b_k f_{t_k}(t).$$

where $f_{t_k}(t)$ is the footprint related to the Dirac at location $t_k$. Assume that we observe the finite resolution version $x_{J_0+1}(t)$. The representation of $x_{J_0+1}$ in terms of footprints is given by

$$x_{J_0+1}(t) = \sum_{n=-\infty}^{\infty} y_{J,n}\varphi_{J,n}(t) + \sum_{k=0}^{K-1} b_k \hat{f}_{t_k}(t)$$

with $\hat{f}_{t_k}$ representing the finite resolution version of $f_{t_k}$. If $x(t)$ satisfies the hypotheses of propositions in Section 3, then we can reconstruct the infinite resolution version of $\hat{f}_{t_k}(t)$ by comparing it with all the possible finite resolution footprints $\hat{f}_{t_x}(t)$ at arbitrary location $t_x$ and by choosing the one that maximizes $\langle \hat{f}_{t_k}, \hat{f}_{t_x}\rangle$. More precisely, assume that $t_x$ is close enough to $t_k$, then

$$\langle x_{J_0+1}(t), \hat{f}_{t_x}(t)\rangle = b_k\langle \hat{f}_{t_k}, \hat{f}_{t_x}\rangle$$

and, it is possible to show that the maximum of $\langle \hat{f}_{t_k}, \hat{f}_{t_x}\rangle$ is achieved only when $t_x = t_k$.

In practice, it is not feasible to compute all the possible inner products $\langle \hat{f}_{t_k}, \hat{f}_{t_x}\rangle$ since $t_x$ is real. However, if one is only interested in enhancing the resolution of $x_{J_0+1}(t)$, then one has to test only a limited number of footprints. Assume, for instance, that the new resolution one wants to achieve is $2^{J_1}$ with $J_1 < J_0$, then the footprints that we need to consider are only at discrete locations $t_n = n \cdot 2^{J_1}$ with $n \in \mathbb{Z}$ and the footprint $\hat{f}_{t_n}(t)$ closest to the actual value $t_k$ gives the highest inner product $\langle \hat{f}_{t_k}, \hat{f}_{t_n}\rangle$ .

An example of the algorithm is illustrated in Figure 2. We consider a periodic piecewise linear signal with period $\tau = 128$ (Figure 2(a)). The signal is sampled with a Daubechies filter with two vanishing moments. The coarse approximation of the signal (what we have called $x_{J_0+1}(t)$) is shown in Figure 2(b). In this case $J_0 = 4$. The reconstruction with footprints of $x(t)$ is shown in Figure 2(c) and is exact to machine precision.[2]

## 5. CONCLUSIONS

In this paper, we have shown that it is possible to sample some classes of signals using a wavelet sampling kernel. We have then developed a new resolution enhancement algorithm based on these sampling results and on the notion of footprints. Future research will focus on the generalization of these sampling theorems to the case of two-dimensional signals and on the design of new algorithms for image resolution enhancement.

[2] Preliminary results seem also to indicate that this algorithm is quite resilient to noise. A more precise analysis of this resilience is under investigation.

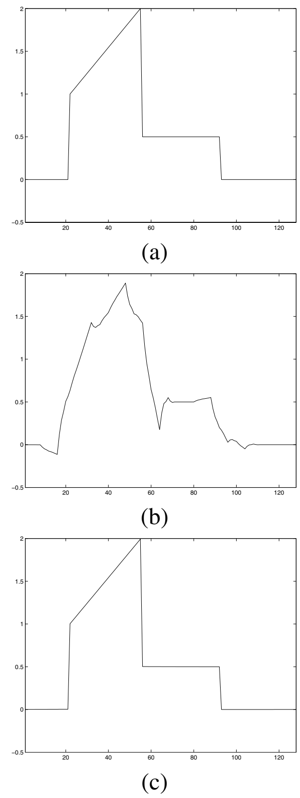


**Fig. 2**. Illustration of the reconstruction algorithm based on footprints. (a) Original discrete-time piecewise linear signal. In this case the original signal has 128 samples. (b) Coarse version of the signal using Daubechies filters with two vanishing moments. This coarse version is obtained taking only 16 samples. This means $J_0 = 4$. (c) Reconstruction with footprints of the original signal using the 16 samples of the coarse version.

## 6. REFERENCES


[1] I. Daubechies. *Ten Lectures on Wavelets*. Society for Industrial and Applied Mathematics, Philadelphia, PA, 1992.

[2] P.L. Dragotti and M. Vetterli. Wavelet footprints: Theory, algorithms and applications. *IEEE Trans. on Signal Processing*, 51(5):1306–1323, May 2003.

[3] S. Mallat. *A Wavelet Tour of Signal Processing*. Academic Press, 1998.

[4] G. Strang and T. Nguyen. *Wavelets and Filter Banks*. Wellesley-Cambridge Press, Boston, 1996.

[5] M. Unser. Sampling-50 years after Shannon. *Proc. IEEE*, 88:569–587, April 2000.

[6] M. Unser and A. Aldroubi. A general sampling theory for nonideal acquisition devices. *IEEE Trans. Signal Processing*, 42(11):2915–2925, November 1994.

[7] M. Vetterli and J. Kovačević. *Wavelets and Subband Coding*. Prentice-Hall, Englewood Cliffs, NJ, 1995.

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