

# Resolution enhancement and sampling with wavelets and footprints

Pier Luigi Dragotti<sup>1</sup> and Martin Vetterli<sup>2,3</sup>

<sup>1</sup> Electrical and Electronic Engineering  
Imperial College, Exhibition Road,  
London SW7 2BT, United Kingdom

<sup>2</sup> École Polytechnique Fédérale de Lausanne (EPFL)  
CH - 1015 Lausanne, Switzerland.

<sup>3</sup> Electrical Engineering and Computer Science Department,  
University of California at Berkeley, Berkeley CA 94720

## 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.<sup>2</sup> Preliminary results show that this algorithm is also very resilient to noise.

## 1. INTRODUCTION

A critical element in modern signal processing and communication is sampling. Often one has no direct access to the original continuous-time phenomenon and can only observe and analyze it through sampling. In some cases the original 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

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

Given the samples  $y_n$ , the key problem is to understand under which conditions one can reconstruct  $x(t)$  from the  $y_n$ 's and which is the best way to perform this reconstruction.

If  $x(t)$  is bandlimited, 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 bandlimited.<sup>9</sup> 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<sup>9</sup> 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 and for robust signal reconstruction in the case the samples  $y_n$  have been corrupted by noise. This algorithm is based on the notion of wavelet footprints which was introduced in.<sup>2</sup> For an excellent review on sampling we refer to.<sup>5</sup> Some pioneering works on sampling with the wavelet transform can be found in.<sup>6</sup>

The paper is organized as follows: The next section presents a review of the wavelet transform and of wavelet footprints. In Section 3, we present new sampling theorems for signal with a finite rate of innovation. Section 4 provides an interpretation of these sampling results in terms of resolution enhancement and in Section 5, we present a footprint-based algorithm for resolution enhancement. We conclude in Section 6.

## 2. WAVELETS AND FOOTPRINTS

This section presents a brief review of the wavelet transform and of wavelet footprints. For a more detailed treatment of wavelets, we refer the reader to.<sup>1, 3, 4, 8</sup> Wavelet footprints were introduced in.<sup>2</sup>

### 2.1. The wavelet transform

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

$$\int_{-\infty}^{\infty} |\Psi(\omega)|^2 |\omega|^{-1} d\omega < \infty,$$

where  $\Psi(\omega)$  is the Fourier transform of  $\psi(t)$  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

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

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

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

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

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

with the usual assumption 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.

A function  $\varphi(t)$  is an admissible scaling function of  $L_2(\mathbb{R})$  if and only if it satisfies the following three conditions<sup>5, 7</sup>:

1. Riesz basis criterion

$$A \leq \sum_{n \in \mathbb{Z}} |\hat{\varphi}(\omega + 2\pi n)|^2 \leq B$$

2. Refinability

$$\varphi\left(\frac{x}{2}\right) = 2 \sum_{k \in \mathbb{Z}} h(k) \varphi(x - k)$$

3. Partition of unity

$$\sum_{k \in \mathbb{Z}} \varphi(t - k) = 1.$$

These conditions ensure that there exists a corresponding wavelet  $\psi(t)$  that generates a basis of  $L_2(\mathbb{R})$ . The wavelet is then expressed as a linear combination of shifted versions of  $\varphi(t)$ :

$$\psi\left(\frac{t}{2}\right) = 2 \sum_{n \in \mathbb{Z}} g_n \varphi(t - n).$$

Finally, 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

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

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,

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

In Section 3, 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.

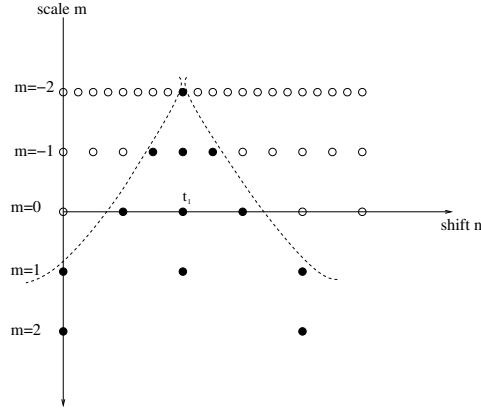
## 2.2. Wavelet footprints

Consider the signal  $x(t) = a_0 \delta(t - t_0)$ , that is, a signal made of a single Dirac at location  $t_0$ , and consider its decomposition in terms of wavelets and scaling functions:

$$x(t) = \sum_{n=-\infty}^{\infty} y_{J,n} \varphi_{J,n}(t) + \sum_{m=-\infty}^J \sum_{n=-\infty}^{\infty} d_{m,n} \psi_{m,n}(t). \quad (6)$$

If we assume that the wavelet  $\psi(t)$  has compact support, then only a limited number of wavelets is influenced by this Dirac, namely, only the wavelets  $\psi_{m,n}(t)$  that overlap the location  $t_0$ . Recall that the set of points  $(m, n)$  such that  $t_0$  is included in the support of  $\psi_{m,n}(t)$  is called cone of influence of  $t_0$  (see Figure 1). Thus, if  $\psi(t)$  has compact support, the only non-zero wavelet coefficients of the decomposition (6) are the coefficients in the cone of influence of  $t_0$ . Moreover, the coefficients inside this cone of influence have only one degree of freedom since their values depend uniquely on the amplitude  $a_0$  of the Dirac. These considerations lead to the definition of wavelet footprints<sup>2</sup>:

**DEFINITION 1.** *Given is a stream of Diracs with only one Dirac at location  $t_0$ , we call footprint  $f_{t_0}$  the scale-space vector obtained by gathering together all the wavelet coefficients in the cone of influence of  $t_0$  and then imposing  $\|f_{t_0}(t)\| = 1$ .*



**Figure 1.** The cone of influence of  $t_1$  consists of the scale-space points  $m, n$  such that  $t_1$  is included in the support of  $\psi_{m,n}(t)$ .

Expressed in terms of wavelets, this footprint can be written as

$$f_{t_0}(t) = \sum_{j=-\infty}^J \sum_{n \in I_0} u_{j,n} \psi_{j,n}(t)$$

where  $I_0$  is the set of indices  $n$  which are in the cone of influence of  $t_0$  and  $u_{j,n}$  are the coefficients of the decomposition (6) normalized to ensure  $\|f_{t_0}(t)\| = 1$ . Clearly, we have a different footprint for every different location  $t_0$  and the scale  $J$  in the footprint definition can be chosen arbitrarily.

Therefore, any signal of the form  $x(t) = a_0 \delta(t - t_0)$  can be expressed in terms of the scaling functions and of  $f_{t_0}(t)$ . For instance, the signal  $x(t)$  in (6) becomes

$$x(t) = \sum_{n=-\infty}^{\infty} y_{J,n} \varphi_{J,n}(t) + b_0 f_{t_0}(t)$$

where  $b_0 = \langle x, f_{t_0} \rangle$ . Likewise, a complete stream of Diracs  $x(t) = \sum_{k \in \mathbb{Z}} a_k \delta(t - t_k)$  can be represented by a linear combination of scaling functions and footprints, that is,

$$x(t) = \sum_{n=-\infty}^{\infty} y_{J,n} \varphi_{J,n}(t) + \sum_{k \in \mathbb{Z}} b_k f_{t_k}(t). \quad (7)$$

The cones of influence of two different footprints at locations  $t_0$  and  $t_1$  may overlap. In this case, we say that the two footprints are biorthogonal, otherwise, we say they are orthogonal. The orthogonality of two footprints depends on the distance between the locations  $t_0$  and  $t_1$ , and on the choice of  $J$ . Assume, for instance, that in the decomposition in (7)  $J$  is chosen so that there are no overlapping footprints, then the coefficients  $b_k$  of that decomposition are given by  $b_k = \langle x, f_{t_k} \rangle$ .

The notion of footprints generalizes to the case of piecewise polynomial signals.<sup>2</sup> First, we need to define a piecewise polynomial signal.

**DEFINITION 2 (VETTERLI, MARZILIANO, BLU<sup>9</sup>).** *A  $\tau$ -periodic signal  $x(t)$  is piecewise polynomial with  $K$  pieces each of maximum degree  $R$  if and only if its  $(R + 1)$  derivative is a stream of differentiated Diracs given by  $x^{(R+1)} = \sum_k \sum_{r=0}^R a_{kr} \delta^{(r)}(t - t_k)$ .*

In a similar way one can define a non-periodic piecewise polynomial signal. Assume that  $x(t)$  is a piecewise polynomial signal, the wavelet decomposition of this signal with a compact support wavelet with at least  $R + 1$

vanishing moments, results in zero wavelet coefficients except for the coefficients in the cones of influence of  $t_k$   $k = 0, 1, \dots, K - 1$ . In this case, one can prove that the wavelet coefficients inside each cone of influence have only  $R + 1$  degrees of freedom and these coefficients can be precisely represented by  $R + 1$  footprints  $f_{t_k}^{(r)}(t)$   $r = 0, 1, \dots, R$ . These footprints are designed to satisfy the following two conditions

$$\|f_{t_k}^{(r)}(t)\| = 1 \quad r = 0, 1, \dots, R.$$

$$\langle f_{t_k}^{(r)}(t), f_{t_k}^{(l)}(t) \rangle = \delta_{r,l}.$$

With this set of footprints one can characterize any discontinuity at a fixed location  $t_k$ . For instance, a piecewise polynomial signals  $x(t)$  with  $K$  discontinuities at locations  $t_k$ ,  $k = 0, 1, 2, \dots, K - 1$  can be written as

$$x(t) = \sum_{n=-\infty}^{\infty} y_{J,n} \varphi_{J,n}(t) + \sum_{k=0}^{K-1} \sum_{r=0}^R b_{k,r} f_{t_k}^{(r)}(t).$$

Finally, assume that  $x(t)$  is a periodic piecewise polynomial signal with period  $\tau$  and assume that  $J = \log_2 \tau$ ;  $x(t)$  can be written as<sup>2</sup>

$$x(t) = y_{J,0} \varphi_{J,0}(t) + \sum_{k=0}^{K-1} \sum_{r=0}^R b_{k,r} f_{t_k}^{(r)}(t) \quad (8)$$

where the coefficient  $y_{J,0}$  is proportional to the mean value of  $x(t)$ . This representation of  $x(t)$  is among the most compact representations of a piecewise polynomial signal, since the number of coefficients in (8) equals the number of degrees of freedom of the piecewise polynomial signal.

In conclusion, footprints represent a simple and efficient way to model the dependency across scales of the wavelet coefficients. Footprints will be used in the rest of the paper to enhance the resolution of signals with a finite rate of innovation.

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

We have seen in the previous section that a valid scaling function has to satisfy three conditions: Riesz basis criterion, refinability and partition of unity. In this section, we further assume that the scaling function is 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 that a linear combination of  $\varphi(t)$  can reproduce polynomials of degree  $K - 1$ .

We consider two classes of signals namely streams of Diracs and piecewise polynomial signals. We present the results for streams of Diracs in details and derive the other sampling theorems directly from these results.

#### 3.1. Streams of Diracs

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

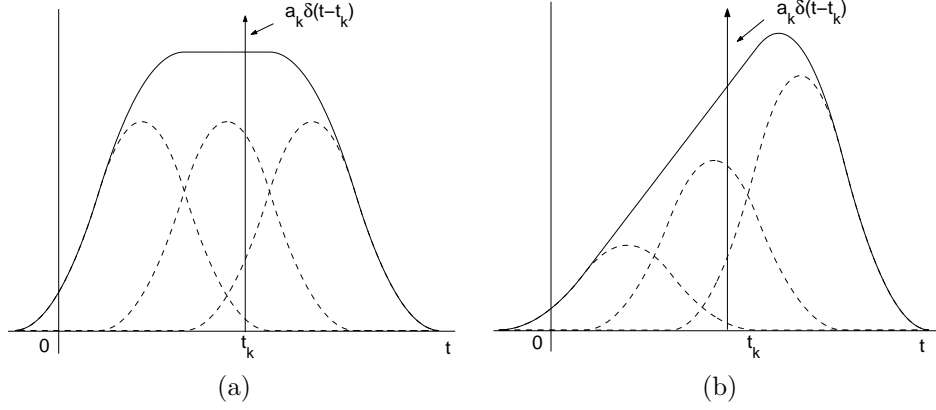
**PROPOSITION 1.** *Given is a scaling function  $\varphi(t)$  that can reproduce polynomials of maximum degree one and of compact support  $L$ . An infinite-length stream of Diracs  $x(t) = \sum_{n \in \mathbb{Z}} a_n \delta(t - t_n)$  is uniquely determined from the samples defined by  $y_n = \langle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

---

\*It is straightforward to show that if there is more than one Dirac in an interval of size  $LT$ , then it is not possible to reconstruct  $x(t)$ , for this reason this part of the proof is omitted.



**Figure 2.** Illustration of the reproduction of polynomials of degree zero and one using a B-spline of degree two. In this case  $L$  is equal to three and only three translated versions of the scaling function overlap the Dirac. 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,  $c_{1,0}y_0 + c_{1,1}y_1 + c_{1,2}y_2 = a_k t_k$  as illustrated in Figure 2(b).

$[n - 3L/2 + 1, -L/2n]$  (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 that  $a_k \delta(t - t_k)$  is. Assume  $y_n$   $n = 0, 1, \dots, L - 1$  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

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

and

$$t_k = (\sum_n c_{1,n} y_n) / a_k \quad (10)$$

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

In equations (9) and (10), we have used the fact that, in the vicinity of  $t_k$ , the scaling function is reproducing polynomials of degree zero and one respectively. In fact, we have that

$$\sum_{n=0}^{L-1} y_n = \langle a_k \delta(t - t_k), \sum_{n=0}^{L-1} \varphi(t - n) \rangle = \int_{-\infty}^{\infty} a_k \delta(t - t_k) (\sum_{n=0}^{L-1} \varphi(t - n)) dt = a_k \sum_{n=0}^{L-1} \varphi(t_k - n) = a_k \quad (11)$$

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

$$\sum_{n=0}^{L-1} c_{1,n} y_n = \langle a_k \delta(t - t_k), \sum_{n=0}^{L-1} c_{1,n} \varphi(t - n) \rangle = a_k \sum_{n=0}^{L-1} c_{1,n} \varphi(t_k - n) = a_k t_k \quad (12)$$

where in the last equality we used the polynomial approximation property (5). Figure 2 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.

In the case of a periodic stream of Diracs we have:

**PROPOSITION 2.** *Given is a scaling function  $\varphi(t)$  that can reproduce polynomials of maximum degree one and of compact support  $L$ , and a periodic stream of Diracs  $x(t)$  of period  $\tau$  with  $K$  Diracs. This signal 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  $LT$  and  $\tau/T \geq LK + 1$ .*

*Proof:* For the hypotheses, there is at least one inner product  $y_n$  which is equal to zero. This is used as an initial condition and then the remaining signal can be determined using the algorithm of Proposition 1.  $\square$

The propositions above have shown conditions under which we can sample streams of Diracs. The reconstruction algorithms rely 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  $LT$ . It is possible to weaken this condition by considering scaling functions that can reproduce higher order polynomials. In that case, one can show that it is possible to reconstruct streams of Diracs with more than one Dirac in an interval of length  $LT$ .

### 3.2. Piecewise Polynomial Signals

We know that by differentiation a piecewise polynomial signal can be reduced to a stream of Diracs (see Definition 2). Thus, we can use the results on streams of Diracs to derive sampling theorems for piecewise polynomial signals.

We know that by using integration by parts we have that

$$\left\langle \frac{d^n x(t)}{dt^n}, y(t) \right\rangle = (-1)^n \left\langle x(t), \frac{d^n y(t)}{dt^n} \right\rangle.$$

For this reason, we can use derivative of the scaling functions and the result in Proposition 1 to sample piecewise polynomial signals.

Although these considerations give us a way to sample piecewise polynomial signals, we prefer to use the scaling function as sampling kernel rather than its  $(R + 1)$  derivative. Therefore, we can enunciate the following result which is weaker, but it will be useful for the rest of the paper.

**PROPOSITION 3.** *A piecewise polynomial signal  $x(t)$  with polynomial pieces of maximum degree  $R$  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  $L$ , if and only if there is at most one discontinuity in an interval of length  $(R + 2)LT$ .*

## 4. INFINITE RESOLUTION WITH FOOTPRINTS

In this section we provide an interpretation of the sampling theorems of Section 3 in terms of resolution enhancement. To simplify the notation we will concentrate on the case of orthogonal wavelets.

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

$$x(t) = \sum_{n=-\infty}^{\infty} y_{J,n} \varphi_{J,n}(t) + \sum_{m=-\infty}^J \sum_{n=-\infty}^{\infty} d_{m,n} \psi_{m,n}(t). \quad (13)$$

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 (13) are sufficient to characterize  $x(t)$  or, in other words, that the finite resolution version  $x_{J+1}(t) = \sum_{n=-\infty}^{\infty} y_{J,n} \varphi_{J,n}(t)$  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 be 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

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

with

$$f_{t_k}(t) = \sum_{j=-\infty}^J \sum_{n \in I_k} u_{j,n} \psi_{j,n}(t).$$

Now, assume that we observe a finite resolution version of  $f_{t_k}$ , that is,

$$\hat{f}_{t_k} = \sum_{j=J_0}^J \sum_{n \in I_k} u_{j,n} \psi_{j,n}(t)$$

with  $J_0 < J$ . The fact that under the hypotheses of the theorems in Section 3, it is possible to reconstruct  $x(t)$  from its finite resolution version, allows us to say that it is in principle possible to increase arbitrarily the resolution of the footprint  $\hat{f}_{t_k}$  and eventually recover the corresponding infinite resolution footprint.

This last observation will be used in the next section to develop a robust resolution enhancement algorithm based on footprints.

## 5. OVERSAMPLING AND RESOLUTION ENHANCEMENT WITH FOOTPRINTS

In this section we investigate the use of footprints to increase the resolution of a sampled signal and to reduce the effect of noise in the case of noisy samples. Again, for the sake of simplicity, we consider orthogonal wavelets and focus on piecewise polynomial signals. In particular, we concentrate on piecewise constant signals.

Consider a piecewise constant signal  $x(t)$ , we have seen that it can be written as

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

or, using footprints, as

$$x(t) = \sum_{n=-\infty}^{\infty} y_{J,n} \varphi_{J,n}(t) + \sum_{k=0}^{K-1} b_k f_{t_k}^{(0)}(t).$$

with  $f_{t_k}^{(0)}(t) = \sum_{j=J_0}^J \sum_{n \in I_k} u_{j,n} \psi_{j,n}(t)$ . Assume that we observe a finite resolution version of  $x(t)$ , namely,

$$x_{J_0+1} = \sum_{n=-\infty}^{\infty} y_{J,n} \varphi_{J,n}(t) + \sum_{m=J_0}^J \sum_{n=-\infty}^{\infty} d_{m,n} \psi_{m,n}(t).$$

This means that  $x(t)$  has been sampled with a scaling function  $\varphi(t/2^{J_0})$ . The representation of  $x_{J_0+1}$  in terms of footprints is given by

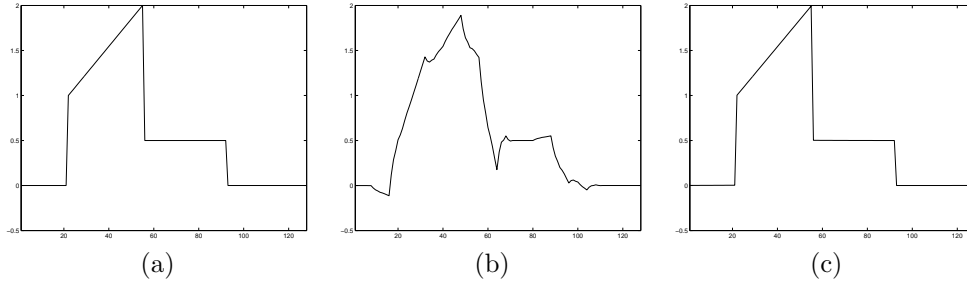
$$x_{J_0+1}(t) = \sum_{n=-\infty}^{\infty} y_{J,n} \varphi_{J,n}(t) + \sum_{k=0}^{K-1} b_k \hat{f}_{t_k}^{(0)}(t),$$

with  $\hat{f}_{t_k}^{(0)}(t) = \sum_{j=J_0}^J \sum_{n \in I_k} u_{j,n} \psi_{j,n}(t)$  representing the finite resolution version of  $f_{t_k}^{(0)}$ . Now, assume that  $x(t)$  satisfies the hypotheses of Proposition 3 and that  $J$  is chosen so that there are no overlapping footprints.<sup>†</sup> We can reconstruct the infinite resolution version of  $f_{t_k}^{(0)}(t)$  by comparing it with all the possible finite resolution

---

<sup>†</sup>Notice that, since we know that there is at most one discontinuity in an interval of length  $2LT$  (Proposition 3), it is always possible to choose  $J$  such that there are no overlapping footprints.





**Figure 3.** Illustration of the reconstruction algorithm based on footprint. (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.

footprints  $\hat{f}_{t_x}^{(0)}(t)$  at arbitrary location  $t_x$  and by choosing the one that maximizes  $\langle \hat{f}_{t_k}^{(0)}, \hat{f}_{t_x}^{(0)} \rangle$ . More precisely, assume that  $t_x$  is close enough to  $t_k$ , the following holds

$$\langle x_{J_0+1}(t), \hat{f}_{t_x}^{(0)}(t) \rangle = b_k \langle \hat{f}_{t_k}^{(0)}, \hat{f}_{t_x}^{(0)} \rangle,$$

and it is possible to show that the maximum of  $\langle \hat{f}_{t_k}^{(0)}, \hat{f}_{t_x}^{(0)} \rangle$  is achieved only when  $t_x = t_k$ .

In practice, it is not feasible to compute all the possible inner products  $\langle \hat{f}_{t_k}^{(0)}, \hat{f}_{t_x}^{(0)} \rangle$  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  $\hat{f}_{t_n}^{(0)}(t)$  closest to the actual value  $t_k$  gives the highest inner product  $\langle \hat{f}_{t_k}^{(0)}, \hat{f}_{t_n}^{(0)} \rangle$ .

Finally, it is worth pointing out that the resolution  $2^{J_0}$  at which we sample  $x(t)$  is arbitrary. In fact,  $J_0$  has to be such that the hypotheses of Proposition 3 are satisfied and must be smaller than  $J$ . If we increase the sampling resolution further, that is, we choose  $J_1 < J_0$ , then the finite resolution footprints that we consider in our algorithm operate at more scales. For this reason the inner products in the algorithm tend to become more robust to noise. Therefore, by increasing the resolution of our sampling, we automatically increase the resilience of our algorithm to noise.

An example of the algorithm is illustrated in Figure 3. We consider a periodic piecewise linear signal with period  $\tau = 128$  (Figure 3(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 3(b). In this case  $J_0 = 4$ . The reconstruction with footprints of  $x(t)$  is shown in Figure 3(c) and is exact to machine precision.

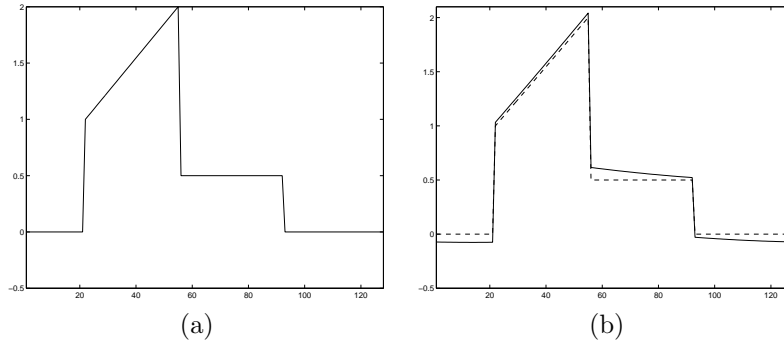
The effect of the noise is shown in Figure 4 and in Table 5. In particular, the table shows how the performance improves with the number of samples.

Samples	8	16	32	64
SNR	20.7dB	22.2dB	23.2dB	24.5dB

**Table 5.** Effect of the oversampling in the reconstruction of a piecewise constant signal. The samples have been corrupted with additive Gaussian noise.

## 6. CONCLUSIONS

In this paper, we have shown that it is possible to sample some classes of signal 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.



**Figure 4.** Effect of the noise in the reconstruction algorithm. In this case, the samples  $y_n$  have been corrupted by additive Gaussian noise. (a) Original signal. (b) Original signal (dashed) and reconstructed signal (SNR=22.9dB).

## 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. Unser and T. Blu. Wavelet theory demystified. *IEEE Trans. Signal Processing*, 51(2):470–483, February 2003.
8. M. Vetterli and J. Kovačević. *Wavelets and Subband Coding*. Prentice-Hall, Englewood Cliffs, NJ, 1995.
9. M. Vetterli, P. Marziliano, and T. Blu. Sampling signals with finite rate of innovation. *IEEE Trans. Signal Processing*, 50(6):1417–1428, June 2002.