Denoising and Harmonic Detection Using Libraries of Nonorthogonal Trigonometric Bases

Mercorelli, Paolo; Terwiesch, Peter; Prattichizzo, Domenico

Published in:
1st International IFAC Symposium on System and Structure and Control

Publication date:
2001

Document Version
Early version, also known as pre-print

Denoising and Harmonic Detection Using Libraries of Nonorthogonal Trigonometric Bases

P. Mercorelli[ab] **P. Terwiesch**[c] **D. Prattichizzo**[a]

[a]Dip. di Ingegneria dell' Informazione – Università di Siena
via Roma, 56, 53100 Siena, Italia
prattichizzo@ing.unisi.it

[b]Automation and Information Technology Department – ABB Corporate Research.
Speyerer Str. 4, D-69115 Heidelberg (Germany).
Tel. +49 6221 596263 – Fax +49 6221 596253
Paolo.Mercorelli@de.abb.com

[c]ABB Corporate Research
Segelhof, CH-5405 Baden (Switzerland)
Tel. +41 56 486 8319 – Fax +41 56 493 5401
Peter.Terwiesch@ch.abb.com

Keywords: Denoising, data compression, trigonometric bases, wavelet packets, rail vehicle control.

1 Abstract

The paper deals with simultaneous noise suppression and signal compression in order to reconstruct and to monitor quasi harmonic signals. In particular one is interested in monitoring and recognizing signals as they occur in rail vehicle systems. One describes an algorithm to estimate discrete signals from their noisy observations using a library of nonorthonormal bases. The described technique combines the *shrinkage technique* and techniques in regression analysis using *Shannon Entropy function* and *Cross Entropy function*. When the problem is to compress and to detect the signals one needs an algorithm which can illuminate the difference between the noise and the desired signal. Recursive residual iterations with cosine and sine bases allow to reconstruct the signal and the noise with the best *discernable* bases. Simulations using real measured data from an electrical railway line are included to illustrate and to analyse the effectiveness of the proposed method.

2 Introduction

The harmonic interference problem in electrical railway systems has received increasing attention in these last years. The widespread utilization of modern electronic devices, as GTO or IGBT, can cause interference in signal circuits and communication systems as well as it can lead to instability problems. Harmonic detection techniques are also of great importance for vehicle distortion current monitoring and in many practical situations not an easy task, their magnitude and phase change over time.

One presents now an algorithm for signal denoising by using libraries of nonorthogonal bases (frames) such as local smooth trigonometric libraries. The method extracts from observed discrete signals a coherent part which is well represented by the given waveforms and a noisy or incoherent part which cannot be *well compressed* by the waveforms. The proposed technique is essentially a nonparametric regression analysis.

The algorithm consists of creating a map for the values of the *Shannon Entropy function* on every time-frequency cell of the sine and cosine packets for the measured signal. One splits the libraries in two classes: a coherent (for instance with cosine bases) and an incoherent (with sine bases) decomposition, by minimizing and by maximizing the Shannon Entropy function and the Cross Entropy function respectively.

One selects the bases with the best compression level and the bases which *illuminate* the difference between the noise (incoherent decomposition) and the signal (coherent decomposition). Recursive residual iterations with biorthogonal cosine bases for the coherent decomposition and with sine bases for the noise allow to reconstruct the signal and the noise with the best *discernable* bases.

It is known that the Shannon entropy function is a measure of the flatness of the energy distribution of the signal so that minimizing this leads to an efficient representation mainly for signal compression [3]. When the problem is to compress and to detect in order to reconstruct the signals one needs an algorithm which can illuminate the difference between the noise and the signals. It is known that the *Cross Entropy function* is a measure of the discrepancy between two or more bases, [7].

Now, why is one using wavelet packets and their relative biorthogonal families in nonorthogonal libraries ? Why smooth trigonometric bases ?

One needs to define a *language* to describe the signals. More, one has to optimize certain criteria depending on the particular problem. The language must be as much versatile as possible and as much elastic as possible in order to describe various local physical features of the signal. In the mean time the method must be computationally efficient to be practically applied. The wavelet frame provides flexible coordinate system, with their *redundant* adaptive time-frequency cells are able to capture the features of the signals in a reasonable computational calculations. The smooth trigonometric bases match very well the desired harmonic signal and they can detect the information in few *coherent data*. More, the nonorthogonal libraries allow more elasticity in order to approximate the measured signals. In fact by relaxing the orthogonality, much more freedom on the choice of the wavelet function is gained to guarantee good choices of the compressed parameters, even though the fast algorithms associated to the orthogonality are lost, [9]. The presented case calls for a good fitting of the data because of the few a priori knowledge. In order to consider and to use the nonorthogonality of the frames which generates an interaction between the elements of the bases[1] the algorithm considers to every step all the elements of the bases previously selected, without any elimination, see [9].

Because of the decomposition on a nonorthogonal basis is not unique one needs to stop the algorithm, for instance, with a threshold criterion at the stage i for the $\mathcal{L}^2$ norm of a differential error.

Simulations using real measured data on the vehicle line are included to illustrate and to analyse the effectiveness of the proposed method.

The paper is organized as follows. In section 3 one defines the problem in an analytical way . In section 4 one discusses briefly several aspects connected to the nonorthogonality and to the smooth trigonometric wavelet packets and the choice of the best regrossor family. In section 5 and in section 6 are devoted to the presentation of the algorithm with its mathematical details.

3 Problem Formulation

Let us consider a discrete degradation model

$$\mathbf{d} = \mathbf{f} + \mathbf{n}, \tag{1}$$

where $\mathbf{d}$, $\mathbf{f}$, $\mathbf{n} \in \mathcal{X} \subseteq \Re^{d_0}$ and $d_0 = 2^{n_0}$. The subspace $\mathcal{X}$ is called *signal space* and d_0 is the number of samples of the signal. The vector $\mathbf{d}$ represents the noisy observed data and $\mathbf{f}$ is the unknown true signal to be estimated. The vector $\mathbf{n}$ is the white Gaussian noise (WGN), its distribution in assumed unknown. One considers an algorithm to estimate $\mathbf{f}$ from the noisy observation $\mathbf{d}$. First, one prepares the bases mentioned in the previous section in particular one selected the trigonometric bases because the signal $\mathbf{f}$ has sinusoidal shape. Let be $\mathcal{B}$ the candidate library packet tree to describe the signal $\mathbf{f}$. For sake of data compression one wants to compress the signal with a small number of parameter k, as small as possible. In the meantime one wants to minimize the distortion between the estimate and the true signal by choosing the most suitable basis $\mathcal{B}_m$. In general taking a large number of k parameters one obtains smaller values of the error but clearly this generates the so called *data overfitting*. Now the conflict is clear and the problem could be performed like a problem of nonparametric model identification with few a priori knowledge.

The problem can be so formulated:

Given a measured data $\mathbf{d} = \mathbf{f} + \mathbf{n}$ *as in (1) where in particular* $\mathbf{f}$ *is the "quasi harmonic" signal,* $\mathbf{n}$ *is a Gaussian noise with a unknown distribution. Given the trigonometric library* $\mathcal{B} = \{\mathcal{B}_1, \mathcal{B}_2, ...\mathcal{B}_M\}$*, let* $\mathbf{f} = \mathbf{W}_m \alpha_m^k$*, where* $\mathbf{W}_m \in \Re^{k \times k}$ *is a nonorthogonal matrix whose column vectors are the basis elements of* $\mathcal{B}_m$ *and the* α_m^k *are the expansion coefficients of* $\mathbf{f}$ *with only k no-zero coefficients. Find a map* $\mathcal{K}$*, called* feature extractor, $\mathcal{K} : \mathcal{X} \rightarrow \mathcal{F} \subset \Re^k$ *(normally more than one) with* $k \ll d_0$ *in order to extract relevant features such that*

$$\min_{\mathbf{W}_m, \alpha_m^k} \hat{\sigma}^2 = \frac{1}{N} \|\mathbf{d} - \mathbf{W}_m \alpha_m^k\|^2. \tag{2}$$

The problem can be attacked in two steps:

Step 1: *Data shrinkage* in order to reduce the number d_0 of samples.

Step 2: Choice of the best subspaces in order to have the time frequency cell basis coordinates.

4 Giving Up on Orthogonality

Wavelet transform and wavelet series are becoming popular in signal processing and numerical analysis. Loosely speaking, a function $f(t)$ can be decomposed into

$$f(t) = \sum_j \sum_n w_{j,n} \psi_{j,n}(t) \tag{3}$$

where the $\psi_{j,n}(t)$ are the wavelet functions, normally obtained by dilating and translating a mother function $\psi(t)$, the index j and n denote the dilation and translation respectively and $w_{j,n}$ is the weight coefficient for $\psi_{j,n}(t)$. The most popular algorithms are related to the orthonormal wavelet bases, see [4], characterized from fast and elegant algorithms. There are, besides these, less used, the *wavelet frames*, see [4], for which the computations of the coefficients are more complicated but which have certain advantages. As wavelet frames consist of nonorthogonal wavelet families, they are *redundant bases*. To be more formal:

Definition 1 *A family of functions* $\{\psi_{j,n}(t); (j,n) \in \mathsf{Z}, t \in \Re\}$ *in a Hilbert space* $\mathcal{H}$ *is called a frame of* $\mathcal{H}$ *if for every element* $f(t) \in \mathcal{H}$ *there are two positive constants* $\mathbf{A}$ *and* $\mathbf{B}$ *such that:*

$$\mathbf{A}\|f(t)\|^2 \leq \sum_{j,n} \|\langle f(t), \psi_{j,n}(t)\rangle\|^2 \leq \mathbf{B}\|f(t)\|^2. \tag{4}$$

[1] In a frame the decomposition is not unique.

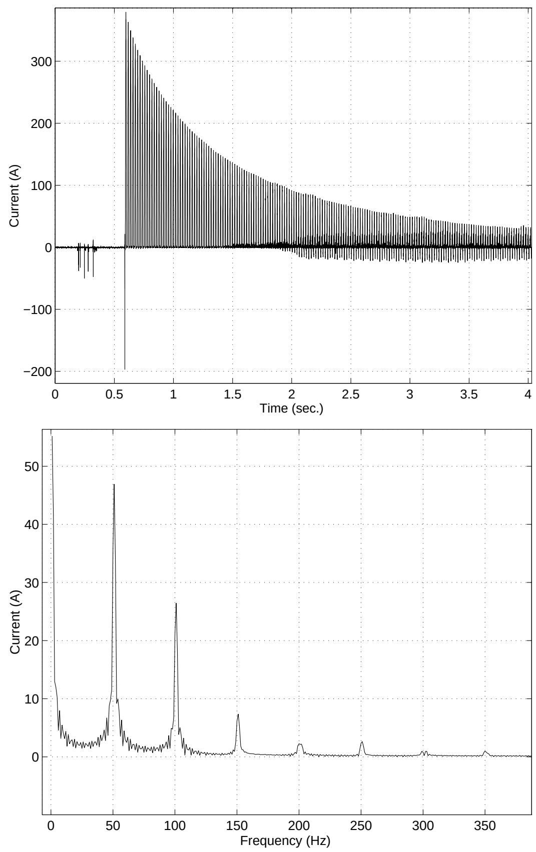

Figure 1: In Top: Inrush current: time domain. In Bottom: Inrush current: windowed spectral analysis.

Where with $\langle \cdot, \cdot \rangle$ one is indicated the inner product and with $\| \cdot \|$ the norm.

In this approach one has a drawback, the optimal decomposition on a nonorthogonal basis is a NP-complete problem and one needs to stop the algorithm, for instance, with a threshold criterion at the stage i for the $\mathcal{L}^2$ norm of the differential error.

Now, the first question is which function to use like activation function. This is a difficult decision, the collected experience on this sense doesn't help too much. All of the model structures are capable of approximating any 'reasonable function' [4]. Thus the question is pick one that 'suits the application', in the sense that only few terms will be needed. A suitable criterion very known in literature is to select the basis which, once fixed a threshold level, has the minimum number of elements in the selected frame. Now, having chosen the best family, how to choose the size of the frame subset ? Finally, how is it possible to select the terms of the subset ?

4.1 Choosing the Best Family Regressor

The case presented in this paper has quasi-harmonic signals that change amplitude and phase over time. This latter aspect suggests the wavelet like activation function. One shows in Fig. 1, where are depicted a measured signal in time domain and its windowed Fourier transform respectively, how the data are very well concentrated around several frequencies, in this case they are multiple of the fundamental (50 Hz). The picture in Fig. 1 seems to suggest a function with a frequency window and time support. A suitable family for this case is the smooth trigonometric wavelet packets. One wants just to recall several basic aspects, further details in [1].

Definition 2 *Let a library of wavelet packets be the collection of functions of the form*

$$\psi_{(d,j,n)}(t) = \psi_j(2^d t - n) \tag{5}$$

where $(d, n) \in \mathsf{Z}$ *and* $j \in \mathsf{N}$.

One is already remarked that one is talking about truncated indices, thus finite libraries of wavelet packets. Here, the *pyramidal* packet is represented with the indices (d, j, n), d is the level of the tree (scaling parameter), j is the frequency cell (oscillation parameter) and n the time cell (localization parameter).

One considers a cover of the real axis $\Re = \bigcup_{-\infty}^{\infty} \mathcal{I}_i$, where $\mathcal{I}_i = [\alpha_i, \alpha_{i+1})$ and $\alpha_i < \alpha_{i+1}$.

Write $\mathcal{T}_i = \alpha_{i+1} - \alpha_i = |\mathcal{I}_i|$ and let $\mathcal{W}_i(t)$ be a window function supported in $[\alpha_i - \frac{\mathcal{T}_{i-1}}{2}, \alpha_{i+1} + \frac{\mathcal{T}_{i+1}}{2}]$ such that

$$\sum_{-\infty}^{\infty} \mathcal{W}_i^2(t) = 1 \tag{6}$$

and

$$\mathcal{W}_i^2(t) = 1 - \mathcal{W}_i^2(2\alpha_{i+1} - t) \qquad for \quad t \quad near \quad \alpha_{i+1}. \tag{7}$$

These conditions tell how the bell function should be taken in order to ensure the orthogonality of the basis. This shows that choosing a basis consisting of adjacent functions, one obtains an orthonormal basis. On the contrary if one considers bases on different levels of the tree don't form an orthonormal basis

The functions

$$\mathbf{S}_{i,k}(t) = \frac{2}{\sqrt{2\mathcal{T}_i}} \mathcal{W}_i(t) sin[(2k+1)\frac{\pi}{2\mathcal{T}_i}(t - \alpha_i)] \tag{8}$$

form an orthonormal basis of $\mathcal{L}^2(\Re)$ subordinate to the partition $\mathcal{W}_i$. The collection of such bases forms a library of orthonormal bases [1]. One can form a library of orthonormal local cosine bases:

$$\mathbf{C}_{i,k}(t) = \frac{2}{\sqrt{2\mathcal{T}_i}} \mathcal{W}_i(t) cos[(2k+1)\frac{\pi}{2\mathcal{T}_i}(t - \alpha_i)]. \tag{9}$$

One has to remark that taking equal smooth windows $\mathcal{W}_i(t)$ (see [1]) then sine/cosine orthogonality can be maintained, see Fig. 2 where one is depicted the sine/cosine bases related to the second level of the packet tree with a frequency of 50 Hz.

It is easy to check that if $\mathbf{H}_i$ denotes the space of functions spanned by $\mathbf{S}_{i,k}$ for $k = 0, 1, 2, ...$ then $\mathbf{H}_i + \mathbf{H}_{i+1}$ is spanned by

$$\mathbf{S}_{i,k}(t) = \mathcal{P}(t) sin[(2k+1)\frac{\pi}{2(\mathcal{T}_i + \mathcal{T}_{i+1})}(t - \alpha_i)], \tag{10}$$

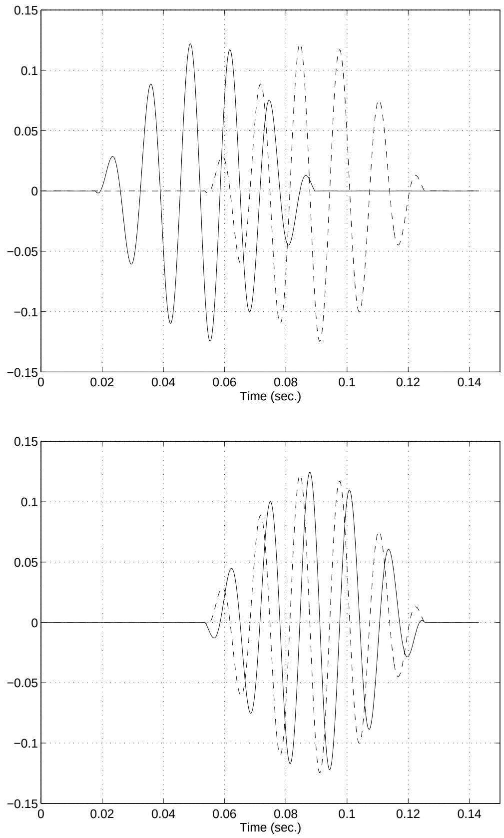

Figure 2: In Top: Biorthogonal smooth local sine and cosine function $\mathbf{S}_{(2,2)}(t)$ and $\mathbf{C}_{(2,2)}(t)$. In Bottom: Adjacent (orthogonal) cosine waveforms with smooth window $\mathbf{C}_{(2,1)}(t)$ and $\mathbf{C}_{(2,2)}(t)$.

where

$$\mathcal{P}^2(t) = \frac{1}{\sqrt{2(\mathcal{T}_i + \mathcal{T}_{i+1})}}\big(\mathcal{W}_i^2(t) + \mathcal{W}_{i+1}^2(t)\big)$$

is a 'window' function covering the interval $I_i \cup I_{i+1}$.

In Fig. 2 one is depicted the sine and cosine biorthogonal functions in the time domain, where one is built the functions with the prototype cutoff $\mathcal{W}(t)$ like a sine function.

It easy to see how the time and frequency cells are linked in a dyadic way, this sort of analysis is equivalent to wavelet packet analysis which allows us to perform an adapted Fourier windowing directly in the time domain. The wavelet packet library is constructed by iterating the wavelet algorithm.

From now on and in order to formally define the *library of wavelet packets* one will consider a new index notation.

Definition 3 *Let a library of wavelet packets be the collection of functions of the form*

$$\psi_{(d,j,n)}(t) = \psi_j(2^d t - n) \tag{11}$$

where $(d,n) \in \mathsf{Z}$ *and* $j \in \mathsf{N}$.

One is already remarked that one is talking about truncated indices, thus finite libraries of wavelet packets. Here, the *pyramidal* packet is represented with the indices (d,j,n), d is the level of the tree (scaling parameter), j is the frequency cell (oscillation parameter) and n the time cell (localization parameter).

The function $\psi_{(d,j,n)}(t) = \psi_j(2^d t - n)$ is roughly centered at $2^{-d}n$, has support of size $\approx 2^{-d}$ and oscillates $\approx j$.

Taking a basis with cells on different level of the tree one obtains a nonorthogonal basis (frames): the symmetry of the windows is lost but not their derivability, they are sums of the derivable functions. In the other words, taking basis elements on different levels of the tree which cover the real axis $\Re$ one is considering superpositions of bases with different resolution frequency cells, the orthogonality is lost. Our algorithm will work transversally on the wavelet packet tree without any restriction in order to use all the possible combinations of the bases, all the possible frames. Once selected the family regressor, for instance the truncated sine/cosine wavelets, the (d,j,n) parameterized family:

$$\Big\{\mathcal{R}_c, \mathcal{R}_s\Big\} = \Big\{\psi^c_{(d,j,n)}(t), \psi^s_{(d,j,n)}(t); (d,n) \in \mathsf{Z}, j \in \mathsf{N}, t \in \Re\Big\}$$

should contain a finite number of wavelets, as less as possible, so that the regressor selection procedure can be efficiently applied. Given an approximating wavelet library not all the wavelet functions are useful, normally only a small number of the coefficients are important, the other ones can be neglected. In general one can select the candidate library as follows:

$$\Big\{\mathcal{R}_c,\ \mathcal{R}_s\Big\} = \Big\{\psi^c_{(d,j,n)}(t), \psi^s_{(d,j,n)}(t) : (d,j,n) \in \mathcal{I}_1 \cup \mathcal{I}_2 \cup \ldots\ldots \cup \mathcal{I}_K\Big\} \tag{12}$$

with $K = 1, 2 \ldots\ldots L$ and

$$\mathcal{I}_k = \Big\{(d,j,n) : \|\psi^c_{(d,j,n)}\|_p > \epsilon, \|\psi^s_{(d,j,n)}\|_p > \epsilon\Big\}, \tag{13}$$

where ϵ is a chosen small positive number. In this way the 'empty' wavelets are eliminated from the wavelet frame. In other words one is starting from a regular tree packet (library) and one selects only those which their support hit our training data. This method is called by some authors *wavelet shrinkage* [2]. One will show that with very few bases of the local trigonometric functions one can obtain a good function detector.

5 Mathematical Details

Let
$$\mathcal{R}_c = \Big\{\psi^c_{(d,j,n)}(t); (d,n) \in \mathsf{Z}, j \in \mathsf{N}, t \in \Re\Big\}$$

and

$$\mathcal{R}_s = \Big\{\psi^s_{(d,j,n)}(t); (d,n) \in \mathsf{Z}, j \in \mathsf{N}, t \in \Re\Big\}$$

be the truncated cosine and sine packet frames respectively as defined in [8], and $\mathbf{Y}_0(t)$ the real row signal.

0. Define the initial condition $\gamma_{c(0)}(t) = \gamma_{s(0)}(t) = \mathbf{Y}_0(t) = \mathbf{d}$ and $\mathbf{f}_{c_0}(t) = 0$, $\mathbf{f}_{s_0}(t) = 0$. Fixed an M index then

Begin-loop

1. For $i = 1, 2, ..M$.
Calculate the weights $\mathbf{c}_{(d,j,n)}$, $\mathbf{s}_{(d,j,n)}$ on all cosine and sine wavelet packet trees according the index:

$$\mathcal{J}\Big(\mathbf{c}_{(d,j,n)}, \mathbf{s}_{(d,j,n)}\Big) = \Big(\mathbf{Y}_{(i-1)}(t) - \sum_{(d,j,n)\in\mathcal{R}_c} \mathbf{c}_{(d,j,n)}\psi^c_{(d,j,n)}(t) - \sum_{(d,j,n)\in\mathcal{R}_s} \mathbf{s}_{(d,j,n)}\psi^s_{(d,j,n)}(t)\Big)^2, \tag{14}$$

this yields:

$$\mathbf{c}_{(d,j,n)} = \Big(\frac{\sum_{(d,j,n)\in\mathcal{R}_c}\langle\gamma_{c(i-1)}(t), \psi^c_{(d,j,n)}(t)\rangle}{\Big(\sum_{(d,j,n)\in\mathcal{R}_c}\big(\psi^c_{(d,j,n)}(t)\big)^2\Big)^{(-1)}}\Big), \tag{15}$$

$$\mathbf{s}_{(d,j,n)} = \Big(\frac{\sum_{(d,j,n)\in\mathcal{R}_s}\langle\gamma_{s(i-1)}(t), \psi^s_{(d,j,n)}(t)\rangle}{\Big(\sum_{(d,j,n)\in\mathcal{R}_s}\big(\psi^s_{(d,j,n)}(t)\big)^2\Big)^{(-1)}}\Big), \tag{16}$$

where $\gamma_{c_{(i-1)}}(t)$ and $\gamma_{s_{(i-1)}}(t)$ are the residual signal of the stage $(i-1)$.

2. Let

$$\mathcal{V}_c = -\sum_{(d,j,n)} \Big(\frac{\hat{\mathcal{P}}\big(\gamma_{c(i-1)}(t)}{\mathcal{P}\big(\gamma_{c(i-1)}(t)\big)}\Big) \ln\Big(\frac{\hat{\mathcal{P}}\big(\gamma_{c(i-1)}(t)\big)}{\mathcal{P}\big(\gamma_{c(i-1)}(t)\big)}\Big), \tag{17}$$

$$\mathcal{V}_s = -\sum_{(d,j,n)} \Big(\frac{\hat{\mathcal{P}}\big(\gamma_{s(i-1)}(t)\big)}{\mathcal{P}\big(\gamma_{s(i-1)}(t)\big)}\Big) \ln\Big(\frac{\hat{\mathcal{P}}\big(\gamma_{s(i-1)}(t)\big)}{\mathcal{P}\big(\gamma_{s(i-1)}(t)\big)}\Big). \tag{18}$$

$$\arg(\min_{\{\mathcal{R}_c\}}(\|\mathcal{V}_c\|)) = \{l_{c_{(d,j,n)}}\}$$

and

$$\arg(\max_{\{\mathcal{R}_s\}}(\|\mathcal{V}_s\|)) = \{l_{s_{(d,j,n)}}\}$$

with $(d,j,n) \in \{\mathcal{R}_s\}$ and with $(d,j,n) \in \{\mathcal{R}_c\}$.

3. Let

$$\mathcal{V}_{(cross)} = \sum_{(d,j,n)\in\mathcal{R}_c}\sum_{(d,j,n)\in\mathcal{R}_s} \Big(\frac{\hat{\mathcal{P}}\big(\gamma_{c(i-1)}(t)\big)}{\mathcal{P}\big(\gamma_{c(i-1)}(t)\big)}\Big) \ln\Big(\frac{\frac{\hat{\mathcal{P}}\big(\gamma_{c(i-1)}(t)\big)}{\mathcal{P}\big(\gamma_{c(i-1)}(t)\big)}}{\frac{\hat{\mathcal{P}}\big(\gamma_{s(i-1)}(t)\big)}{\mathcal{P}\big(\gamma_{s(i-1)}(t)\big)}}\Big),$$

$$\arg(\max_{\{\mathcal{R}_c,\mathcal{R}_s\}}(\mathcal{V}_{(cross)})) = \{l_{c_{(d,j,n)}}, l_{s_{(d,j,n)}}\}$$

with the true probability $\mathcal{P}\big(\gamma_{c(i-1)}(t)\big) = \|\gamma_{c(i-1)}(t)\|^2$ and the estimated probability is calcuated as

$$\hat{\mathcal{P}}\big(\gamma_{c(i-1)}(t)\big) = \|\sum_{(d,j,n)\in\mathcal{R}_c}\mathbf{c}_{(d,j,n)}(t)\psi^c_{(d,j,n)}(t)\|^2.$$

In the same way $\mathcal{P}\big(\gamma_{s(i-1)}(t)\big) = \|\gamma_{s(i-1)}(t)\|^2$ and the

$$\hat{\mathcal{P}}\big(\gamma_{s(i-1)}(t)\big) = \|\sum_{(d,j,n)\in\mathcal{R}_s}\mathbf{s}_{(d,j,n)}(t)\psi^s_{(d,j,n)}(t)\|^2.$$

4. Update $\mathbf{f}_c(t)$, $\gamma_c(t)$, $\mathbf{f}_s(t)$ and γ_s(t):

$$\mathbf{f}_{c_i}(t) = \mathbf{f}_{c_{(i-1)}}(t) + \sum_{l_c\in\mathcal{R}_c}\mathbf{c}_{l_{c_{(d,j,n)}}}\psi^c_{l_{c_{(d,j,n)}}}(t);$$

$$\mathbf{f}_{s_i(t)} = \mathbf{f}_{s_{(i-1)}}(t) + \sum_{l_{s_{(d,j,n)}}\in\mathcal{R}_s}\mathbf{s}_{l_{s_{(d,j,n)}}}\psi^s_{l_{s_{(d,j,n)}}}(t);$$

$$\gamma_{ci}(t) = \gamma_{c(i-1)}(t) - \sum_{l_{c_{(d,j,n)}}\in\mathcal{R}_c}\mathbf{c}_{l_{c_{(d,j,n)}}}(t)\psi^c_{l_{(d,j,n)}}(t);$$

$$\gamma_{si}(t) = \gamma_{s(i-1)}(t) - \sum_{l_{s_{(d,j,n)}}\in\mathcal{R}_s}\mathbf{s}_{l_{s_{(d,j,n)}}}(t)\psi^s_{ls_{(d,j,n)}}(t).$$

$$\mathbf{Y}_i(t) = \mathbf{Y}_{(i-1)}(t) - \mathbf{f}_{c_i}(t) - \mathbf{f}_{s_i}(t)$$

End Loop.

6 Simulations and Results

The case which we present consists of a harmonic signal as presented in the previous sections. As said, this case justifies the choice of the trigonometric bases in order to perform a coherent and an incoherent expansion. We simulated a white Gaussian noise in superposition performed from 0 Hz to 250 Hz. How it is possible to see from the Fig. 3 we can reconstruct the original signal with a good precision. The considered dyadic signal belongs to the space $\Re^{512}$, after the *data shrinkage* the compression is performed and we obtain a dyadic vector belonging to the space $\Re^8$. In other words we can just consider 8 frequencies [0, 50, 100, 150, 200, 250, 300, 350] Hz for every time frequency cell. The selected wavelet packet tree has three levels and considering the Nyquist frequency of data equal to 3500 Hz and the length of the basis equal to 512 samples we obtain a resolution around 3.5 Hz, 7 Hz and 14 Hz respectively[2].

[2]One recalls that the resolution $R = \frac{N_f 2^l}{N}$, where N is the length of the basis, l the level of the tree and N_f the Nyquist frequency

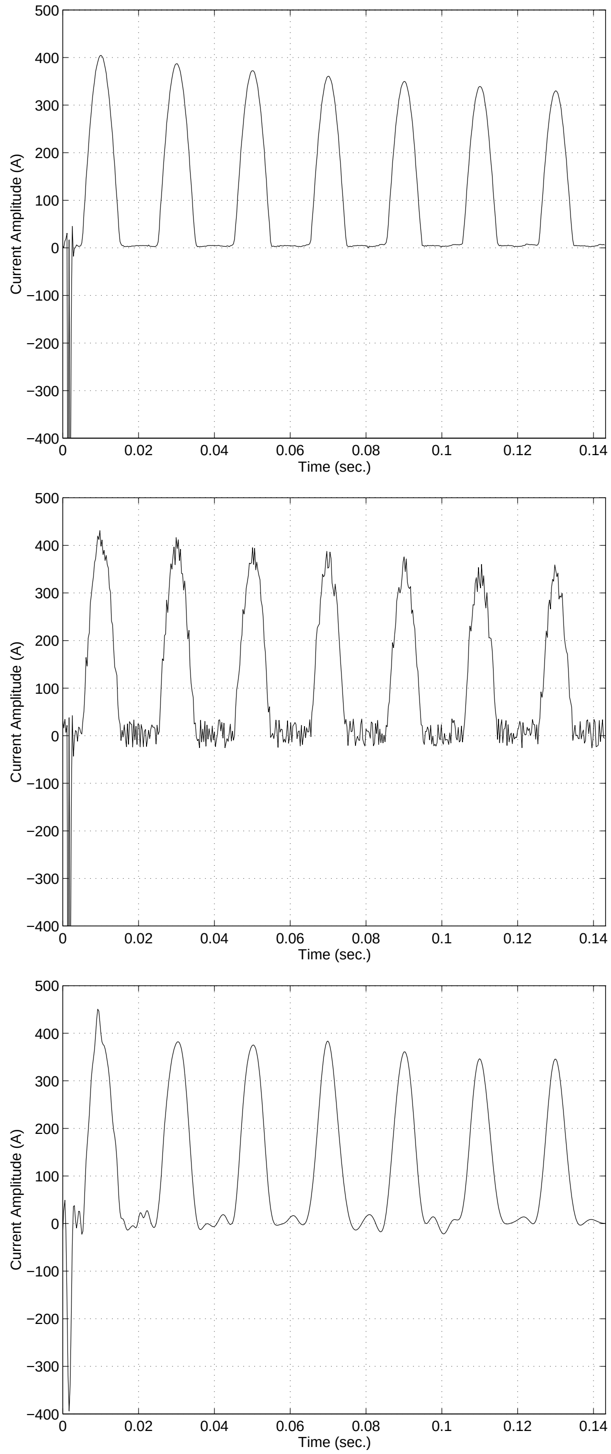

Figure 3: Real Signal. From the Top: Inrush current. Inrush current with noise in superposition. Inrush current through the denoising algorithm.

7 Conclusions

We have described an algorithm for simultaneously suppressing the additive white Gaussian noise component and compressing the signal component in a data set. The bases in the library consist of wavelets, more precisely they consist of wavelet packets where the functions are local trigonometric bases. Cosine and sine bases with their biorthogonality allow to perform an efficient system coordinate. The bases are selected during every step by maximizing the cross entropy function which illuminates the difference between the noise and the desired signal. An example is reported to shown the applicability and usefulness of the algorithm.

References

[1] P. Auscher, G. Weiss, and M. V. Wickerhauser. *Wavelet-A Tutorial in Theory and Applications. Local sine and cosine bases of Coifman and Meyer and the construction of smooth wavelets.* C-K Chui (ed.) ISBN 0-12-174590-2, Academic Press, Boston, 1992.

[2] A. Benveniste, A. Juditscky, B. Delyon, Q. Zhang, and P. G. Glorenee. Wavelets in identification. In *Proc. SYSID'94 10th IFAC Symp. Syst. Identification*, Copenhagen, 1994.

[3] R. R Coifman and M. V. Wickerhauser. Entropy based algorithm for best basis selection. *IEEE Trans. Inform. Theory*, 32:712–718, 1992.

[4] I. Daubechies. *Ten Lectures on Wavelets.* Publisher Society for Industrial and Applied Mathematics, Philadelphia (Pennsylvania), 1995.

[5] Y. Meyer. De la recherche pétrolière à la géometrie des espaces de banach en passant par les paraproduits. *Preprint: École Polytechnique, Palaiseau, Seminaire équations aux dérivées partielles*, 1986.

[6] N. Saito. Simultaneous noise suppression and signal compression using a library of orthonormal bases and the minimum description length criterion. *Wavelets in Geophysics. E. Foufoula-Georgiou and P. Kumar (eds), Academic Press*, 1994.

[7] N. Saito and Ronald R. Coifman. Local discriminant bases. *Mathematical Imaging: Wavelet Applications in Signal and Image Processing II, Proc. SPIE*, 2303, 1994.

[8] P. Terwiesch and P. Mercorelli. A signal classification algorithm using smooth local trigonometric bases. *Submitted to IEEE Transactions on Neural Networks. (Available via ftp in www.dii.unisi.it/ mercorel/)*, 1999.

[9] Q. Zhang. Using wavelet network in nonparametric estimation. *IEEE Transaction on Neural Networks*, 8(2):227–236, 1997.