



HAL
open science

Analysis and short-time extrapolation of stock market indexes through projection onto discrete wavelet subspaces

Laurent Gosse

► **To cite this version:**

Laurent Gosse. Analysis and short-time extrapolation of stock market indexes through projection onto discrete wavelet subspaces. *Nonlinear Analysis: Real World Applications*, 2010, 11, pp.3139-3154. 10.1016/j.nonrwa.2009.11.009 . hal-00414210

HAL Id: hal-00414210

<https://hal.science/hal-00414210>

Submitted on 8 Sep 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Analysis and short-time extrapolation of stock market indexes through projection onto discrete wavelet subspaces

Laurent Gosse

*IAC-CNR “Mauro Picone” (sezione di Bari)
Via Amendola 122/I - 70126 Bari, Italy*

Abstract

We consider the problem of short-time extrapolation of blue chips' stocks indexes in the context of wavelet subspaces following the theory proposed by X.-G. Xia and co-workers in a series of papers [29,28,14,15]. The idea is first to approximate the oscillations of the corresponding stock index at some scale by means of the scaling function which is part of a given multi-resolution analysis of $L^2(\mathbb{R})$. Then, since oscillations at a finer scale are discarded, it becomes possible to extend such a signal up to a certain time in the future; the finer the approximation, the shorter this extrapolation interval. At the numerical level, a so-called Generalized Gerchberg-Papoulis (GGP) algorithm is set up which is shown to converge toward the minimum L^2 norm solution of the extrapolation problem. When it comes to implementation, an acceleration by means of a Conjugate Gradient (CG) routine is necessary in order to obtain quickly a satisfying accuracy. Several examples are investigated with different international stock market indexes.

Key words: Multi-resolution analysis, wavelet decomposition, scale-limited extrapolation, empirical finance.

1991 MSC: 91B84; 94A12; 42C40; 60G25

1 Introduction

Extrapolation is one of the fundamental problems arising in Signal Processing; during the last decades, a lot of work has been devoted to it, in the particular context of band-limited signals, see *e.g.* [11,21,31,12,18,23,26]. The theoretical basis for band-limited extrapolation has to do with the classical Paley-Wiener theorem which states that a function belonging to $L^2(\mathbb{R})$ whose Fourier transform has compact support can be extended to the complex plane as an entire function of exponential type. Hence, its knowledge restricted to any connected

Email address: l.gosse@area.ba.cnr.it (Laurent Gosse).

open set still allows for a reconstruction up to any arbitrarily big domain by relying onto classical analytic functions theory. Numerically, things are more involved, especially because this analytic extension turns out to be an ill-conditioned problem which is quite sensitive to noise and truncation errors; in particular, methods based on Taylor series have been completely given up for Fast Fourier Transform (FFT) routines, see again [11,21,12,18,26].

In this paper, we follow a similar methodology to produce reliable extrapolations of liquid stock market indexes relying on so-called *scale-limited* approximation. Roughly speaking, in the case of band-limited extrapolation, one assumes that the signal under consideration can be approximated by a function of compact support in the Fourier space (hence having only finite frequency oscillations) whereas in the present context, one considers that the signal can be well represented by its projection onto a subspace of $L^2(\mathbb{R})$ containing functions with rather “thick” oscillations. Such subspaces can be found by considering what is now usually called a *Multi-Resolution Analysis* of $L^2(\mathbb{R})$ (see Definition 1 below): they are spanned by bases of shift-invariant *scaling functions* with various smoothness properties [17]. Hence this work applies primarily to indexes which are dominated by their low-frequency approximation and for which fine scale features can be considered as detail not being of great importance.

This is not a very restrictive assumption since the composition of most of modern stock indexes is regularly updated to keep their variations under control even when some share’s prices may have erratic behavior. For instance, on June 18, 2007, AGF and Thomson have been removed from the French index CAC 40; they have been replaced by Air France-KLM and Unibail. Similarly, the Dow Jones Industrials 30 Index will see its composition modified as General Motors should leave because its capitalization dropped too strongly during March 2009. An estimate of the volatility of an investment can be given by a statistical measure known as the standard deviation of the return rate. Standard deviation can be thought as being synonymous with volatility in such a context. An S&P 500 index fund has a standard deviation of about 15%; a standard deviation of zero would mean a return rate that never varies, like a bank account paying compound interest at a guaranteed rate. Exchange rates on FOREX markets seem to display quite an opposite behavior and the methods presented in this paper don’t seem to apply (see *e.g.* [6] for some empirical differences between stock and FX markets, especially the gain/loss asymmetry).

Recently, a whole theory has been proposed in a series of papers [29,28,14,15] (see also [7]) to extend the well-known band-limited extrapolation theory to a more flexible framework of scale-limited spaces which have the nice property to be able to handle finite-duration signals. The roadmap really follows classical Fourier extrapolation and proceeds through linear integral operator theory (see *e.g.* [34] for a general reference): this is recalled in §2.1. However, one big difference is that from a compact set of observations, one cannot theoretically reconstruct the signal on the whole real line (except in the special case where the multi-resolution analysis involves bases of analytic functions), see Theorem 3. So one can merely hope to reconstruct a small part of the missing signal on both sides of the measures interval; hence the “short-time extrapolation” in the title. The size of this “small part” is related to the smoothness of the analyzing wavelets and the scale at which one decides to work (and this scale itself is related to the volatility of the index).

In §2.2, we present the standard algorithms to produce these extrapolations: they originate from the now classical Gerchberg-Papoulis algorithm [11,21] widely used for band-limited extrapolation and whose convergence has been studied in various papers, see [8,12,18,31] and references therein. In the present case, this routine converges toward the so-called *Minimum-norm Solution* of the extrapolation problem. Because of its poor practical convergence properties, an acceleration by the *Conjugate Gradient* (CG) algorithm is generally needed; this is explained in [12,26] in case of band-limited extrapolation, and in [14] for scale-limited extrapolation. We don’t completely follow this paper in §2.3 as we propose an easier way to produce a CG acceleration.

Our specific application is considered in §3. First, in §3.1, we explain that because of the

features of the market indexes, particular wavelet bases are needed. Namely, since these indexes are strongly non-periodic, one cannot consider usual wavelet bases defined on the whole real line as they process finite duration signals either by *zero-padding* or by artificial periodization. Both techniques will lead to the appearance of strong discontinuities at the edge of the observation interval, which will be reflected in the wavelet coefficients at very fine scales. Hence such bases will yield poor representations at medium scales, especially in the vicinity of the borders. A solution is to use the *wavelet bases on a bounded interval*: these bases have been first proposed by Meyer [16], then by Cohen, Daubechies, Jawerth and Vial [3,5]; see also *e.g.* [4,17,19,13]. Then, in §3.2, we explain through an example that results can be of quite poor quality when scale-limited extrapolation is applied to individual share prices displaying locally strong variations on a fine scale. Smoothness and weak volatility are therefore requested to produce an acceptable result.

In §4, we present some numerical results on real-life data. First, in §4.1, a quiet bull market on the French CAC 40 is considered: this gives the best conditions in which the scale-limited extrapolation can be used because the market follows a stable trend and the volatility is low. Then, in §4.2, we look at the real-estate EPRA Eurozone index as a prototype of a moderately volatile stock index in a bear market; the quality of the extrapolation is lower, but there are signs (like the size of some wavelet coefficients) which give a hint about the difficulty in treating such a situation. Finally, we considered what happened on the Hang-Seng Chinese index during the first days of october 2008 when it lost around 25% of its value in less than 10 days; obviously, extrapolation routines cannot give reliable information in such a case, but the analysis of what is left behind after projecting the index onto a scale-limited subspace of $L^2(\mathbb{R})$ can somewhat reveal itself as a warning sign its high volatility (which makes violent corrections plausible in a difficult economic period). In Appendix A, we briefly present results on scale-limited extrapolation of assets log-returns (see [6]).

Let us recall finally that other works previously addressed the wavelet techniques for stock market data: see the review paper [24] and the book [10]. Concerning this book, the authors works primarily with the discontinuous Haar wavelet which in general forbids to use any of these scale-limited extrapolation techniques. Other forecasting algorithms are proposed in [1,2,9,25,32]. A book is devoted to wavelet techniques for time series analysis, see [22]. Some results of empirical finance are reviewed in [6].

2 From band-limited to scale-limited extrapolation

In this section, we aim at reviewing some of the results from [29,28,14,15] which will be useful in the context of processing stock market indexes. Let's begin by introducing the concept of *Multi-Resolution Analysis* (MRA): (see *e.g.* [17] for details)

Definition 1 *A sequence of nested subspaces V_j is called a Multi-Resolution Analysis of $L^2(\mathbb{R})$ if: $\{0\} \subset \dots \subset V_{-1} \subset V_0 \subset V_1 \subset \dots \subset L^2(\mathbb{R})$. Moreover, the following properties must hold:*

- for all $f \in L^2(\mathbb{R})$, $\|\mathcal{P}_{V_j} f - f\|_{L^2} \rightarrow 0$ as $j \rightarrow +\infty$ also, $\mathcal{P}_{V_j} f \rightarrow 0$ as $j \rightarrow -\infty$.
- if $f(t) \in V_j$, then $f(t/2) \in V_{j-1}$ and for all $k \in \mathbb{Z}$, $f(t - 2^j k) \in V_j$.
- there exists a shift-invariant orthonormal base of V_0 given by the scaling function $\phi_n(t) = \phi(t - n)$ for $n \in \mathbb{Z}$.

In this definition, \mathcal{P}_{V_j} stands for the orthogonal projector onto the subspace V_j . Intuitively, it asks for the V_j 's to be vector subspaces of functions with increasing temporal resolution: when j decreases, functions in V_j tend to become constants. Oppositely, when j increases, they are allowed to oscillate with high instantaneous frequency. The wavelet spaces W_j are defined as the orthogonal complement of V_j inside V_{j+1} , that is to say: for all $j \in \mathbb{Z}$, $V_{j+1} = V_j \oplus W_j$.

From ϕ_n , the base of V_0 , one can deduce a base of V_j by simple dilatation,

$$\phi_{j,n}(t) = \sqrt{2^j} \phi_n(2^j t) = \sqrt{2^j} \phi(2^j t - n). \quad (1)$$

Thus, the orthogonal projection of f onto V_j reads:

$$\mathcal{P}_{V_j} f = \sum_{n \in \mathbb{Z}} \langle f, \phi_{j,n} \rangle \phi_{j,n}, \quad \langle f, \phi_{j,n} \rangle = \int_{\mathbb{R}} f(t) \phi_{j,n}(t) dt, \quad (2)$$

which is the best approximation of f in V_j in the least-squares sense. Similarly, for any value of $T \in \mathbb{R}^+$, we denote by \mathcal{P}_T the time-truncation operator: $\mathcal{P}_T f(t) = f(t)$ if $|t| \leq T$, zero in the other case. Finally, χ_A stands for the indicator function of any set A .

2.1 Reproducing kernel Hilbert spaces (RKHS) and continuous extrapolation

The general problem of extrapolation for scale-limited functions can be presented as follows: we fix a MRA in the sense of Definition 1 and an index $j \in \mathbb{Z}$. Any function $f \in V_j$ can be represented by means of the scaling function through the formula (2); in particular, if we choose a compact time interval $[-T, T]$, with $T \in \mathbb{R}^+$, we can define g , a ‘‘truncated version’’ of f the following way: (notice that in general, $g \notin V_j$)

$$\forall t \in [-T, T], \quad g(t) = (\mathcal{P}_T \circ \mathcal{P}_{V_j} f)(t) = \sum_{n \in \mathbb{Z}} \langle f, \phi_{j,n} \rangle \phi_{j,n}(t). \quad (3)$$

Implicitly, this defines a linear bounded operator $\mathcal{H}_j : V_j \rightarrow L^2(-T, T)$. The problem of extrapolating a time-truncated, scale-limited signal g can be formulated as finding a unique solution $f \in V_j$ to $\mathcal{H}_j f = g$. In order to proceed with this inversion, we recall from linear operator theory the adjoint \mathcal{H}_j^* which we can compute easily:

$$\begin{aligned} \langle \mathcal{H}_j f, \varphi \rangle &= \langle \sum_{n \in \mathbb{Z}} \langle f, \phi_{j,n} \rangle \phi_{j,n}, \varphi \rangle \\ &= \sum_{n \in \mathbb{Z}} \langle f, \phi_{j,n} \rangle \langle \phi_{j,n}, \varphi \rangle \\ &= \langle f, \sum_{n \in \mathbb{Z}} \phi_{j,n} \langle \phi_{j,n}, \varphi \rangle \rangle \\ &= \langle f, \mathcal{H}_j^* \varphi \rangle. \end{aligned}$$

Thus \mathcal{H}_j^* is a projector $L^2(-T, T) \rightarrow V_j$; in particular, the composition $\mathcal{H}_j \circ \mathcal{H}_j^*$ defines an integral operator from $L^2(-T, T)$ onto itself,

$$\forall g \in L^2(-T, T), \quad (\mathcal{H}_j \circ \mathcal{H}_j^* g)(t) = \int_{-T}^T g(s) \mathcal{Q}_j(s, t) ds,$$

where \mathcal{Q}_j is called the *Reproducing Kernel* for the *reproducing kernel Hilbert space* [30] V_j :

$$\mathcal{Q}_j(s, t) = \sum_{n \in \mathbb{Z}} \phi_{j,n}(s) \phi_{j,n}(t) = \mathcal{Q}_j(t, s), \quad (s, t) \in \mathbb{R}^2. \quad (4)$$

In the sequel, we assume \mathcal{Q}_j to be bounded¹ as a function of \mathbb{R}^2 and continuous in $[-T, T]^2$.

Lemma 1 (see [14]) *The following properties hold for \mathcal{Q}_j bounded and continuous:*

¹ This is the case as soon as $|\phi(t)| \leq O(1 + |t|^{0.5+\epsilon})^{-1}$ for some $\epsilon > 0$, see [28]; especially, this is true for scaling functions with compact support.

- if the scaling function ϕ is real-valued, $\mathcal{H}_j \circ \mathcal{H}_j^*$ is self-adjoint and positive-semidefinite; in particular, all its eigenvalues are real and nonnegative, so $+\infty > \lambda_1 \geq \lambda_2 \geq \dots \geq 0$.
- $\mathcal{H}_j \circ \mathcal{H}_j^*$ is compact; in particular, its set of eigenfunctions $r_k(t)$ satisfying $(\mathcal{H}_j \circ \mathcal{H}_j^*)r_k = \lambda_k r_k$ forms an orthogonal basis of $L^2(-T, T)$.

The second part of the lemma is classical: see for instance §4.4 (page 240) of [34]. These eigenfunctions are the key to extrapolate scale-limited signals. In case some eigenvalues vanish, let K stand for the biggest integer for which there holds:

$$k \leq K \Rightarrow \lambda_k > 0 \text{ and } \lambda_{K+1} = 0.$$

We then extend the domain of definition of the eigenfunctions r_k from $[-T, T]$ to \mathbb{R} : recall that $\lambda_k r_k(t) = \int_{-T}^T r_k(s) \mathcal{Q}_j(s, t).ds$. Hence, by definition of the reproducing kernel \mathcal{Q}_j (4), we can build for all $k \leq K$,

$$\forall t \in \mathbb{R}, \quad \tilde{r}_k(t) = \frac{1}{\lambda_k} \sum_{n \in \mathbb{Z}} \left(\int_{-T}^T r_k(s) \phi_{j,n}(s) \right) \phi_{j,n}(t) \in V_j.$$

It turns out that these modified eigenfunctions have interesting properties:

Lemma 2 (see [14]) *For \mathcal{Q}_j bounded and continuous, the set of functions $\tilde{r}_k \in V_j$, $k \leq K$ is orthonormal in $L^2(\mathbb{R})$ and orthogonal in $L^2(-T, T)$; their corresponding eigenvalues are positive and lower than 1. Moreover, there holds:*

$$\forall f \in V_j, \quad \int_{-T}^T f(s) r_k(s).ds = \lambda_k \int_{\mathbb{R}} f(s) \tilde{r}_k(s).ds. \quad (5)$$

In this context, λ_k can be seen as the energy contribution of \tilde{r}_k inside the interval $[-T, T]$; to see this, we choose $f = \tilde{r}_k$ inside the formula (5), apply Cauchy-Schwarz inequality, and use $\|\tilde{r}_k\|_{L^2(\mathbb{R})} \geq \|\tilde{r}_k\|_{L^2(-T, T)}$ which leads to:

$$\|\tilde{r}_k\|_{L^2(\mathbb{R})} \leq \frac{1}{\lambda_k} \|r_k\|_{L^2(-T, T)}.$$

We can now define \mathcal{U}_j , the linear subspace of $L^2(\mathbb{R})$ spanned by the set \tilde{r}_k , $k \leq K$.

Theorem 3 (see [14, 28]) *The following decomposition holds for any $j \in \mathbb{Z}$:*

$$V_j = \mathcal{U}_j \oplus \left(V_j \cap \mathcal{P}_T \left[L^2(\mathbb{R}) \right]^\perp \right), \quad \mathcal{U}_j = \mathcal{P}_{V_j} \circ \mathcal{P}_T \left[L^2(\mathbb{R}) \right].$$

As a consequence of the definition of $\mathcal{U}_j \subset V_j$, any function $f \in V_j$ can be written as $f_1 + f_2$ with $f_1 \in \mathcal{U}_j$ and $f_2 \in \mathcal{U}_j^\perp$. Obviously, $f_1(t) = \sum_{k \leq K} a_k \tilde{r}_k(t)$. The fact is that, from (5), it is readily seen that the coefficients a_k depend on $\mathcal{P}_T f$ only:

$$a_k = \int_{\mathbb{R}} f(s) \tilde{r}_k(s).ds = \frac{1}{\lambda_k} \int_{-T}^T f(s) r_k(s).ds.$$

At this point, there is no clue to compute the missing part f_2 which belongs to the set of functions in V_j vanishing on $[-T, T]$. However, we can get rid of this part by a smoothness argument if we aim at extrapolating the signal on a small enough interval. Indeed, for finite j , the space $V_j \cap L^2(-T, T)^\perp$ cannot contain discontinuous functions; hence there is a small interval beyond T , say $[T, T + \delta T]$, inside which f_2 must vanish. The value of δT

is not known precisely, but clearly, δT is lower than the support of $\phi_{j,0}$. Hence, in case the following rough estimate holds,

$$|\delta T| < |\text{supp}(\phi_{j,0})|, \quad (6)$$

Theorem 3 ensures that a signal $g = \mathcal{P}_T f$, $f \in V_j$, can be extrapolated uniquely up to $[-T - \delta T, T + \delta T]$. The value of δT thus decreases when j is getting bigger, hence a rule of thumb could be stated like “**the more details one decides to keep inside the signal, the shorter the extrapolation length will be**”. It can happen that stock indexes, during periods of high volatility, ask for a big value of j in order to keep track of some of the small scale features: in such a case, the length of the extrapolation interval given by δT will be reduced accordingly through (6) (see an illustration in §4).

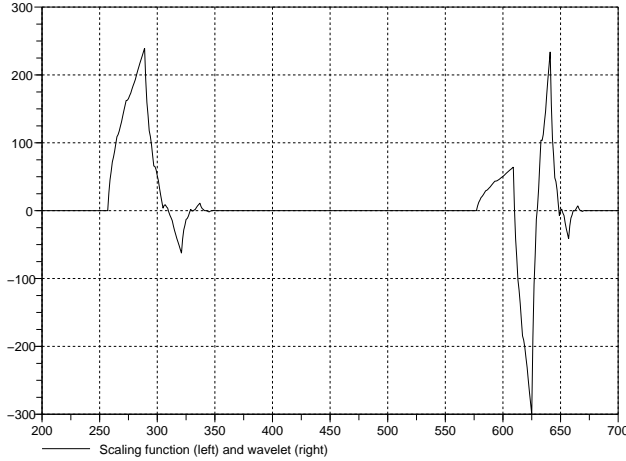


Fig. 1. Compact support of a scaling function and its mother wavelet.

Remark 1 In case the scaling function is analytic, the space $V_j \cap L^2(-T, T)^\perp$ is reduced to zero because a non-zero analytic function cannot vanish on a non-trivial interval. In this case, we recover the usual framework of band-limited extrapolation (see [28,26]).

2.2 Generalized Gerchberg-Papoulis (GGP) algorithm and scale-limited extrapolation

The leading idea of the series of papers [29,28,14,15] is to extend the classical Gerchberg-Papoulis algorithm, originally proposed for band-limited extrapolation (see for instance [11,21,31,18,12,8,26]), to the processing of scale-limited functions belonging to wavelet subspaces. To begin with, let's recall the original construction: a function $f \in L^2(\mathbb{R})$ is said to be *W-band-limited* if its Fourier transform \hat{f} has compact support. In such a case, one has:

$$\forall t \in \mathbb{R}, \quad f(t) = \left(\mathcal{F}^{-1}(\hat{f}) \right)(t) := \int_{-W}^W \hat{f}(\xi) \exp(2i\pi t\xi) . d\xi \quad \text{with } \mathcal{F} : f \mapsto \hat{f}.$$

The number $W \in \mathbb{R}^+$ is called the *band-width* of f . Moreover, the classical Paley-Wiener theorem ensures that f is actually an entire function of exponential type which can be extended to the whole complex plane as a result of analytic continuation theory. This paves the way to various extrapolation routines as we explain now; first, we denote by \mathcal{P}_W the projector onto the subspace of W-band-limited functions. More precisely,

$$\forall f \in L^2(\mathbb{R}), \quad \mathcal{P}_W(f) = \mathcal{F}^{-1} \left(\hat{f} \chi_{[-W,W]} \right).$$

By analogy with the preceding section, we define g as a time-truncated version of a W -band-limited signal f as follows:

$$\forall t \in [-T, T], \quad g(t) = \left(\mathcal{P}_T \circ \mathcal{P}_W(f) \right)(t).$$

Similarly, this defines a bounded linear operator \mathcal{H}_W from the subspace of W -band-limited functions onto its restriction to functions defined in $[-T, T]$. The extrapolation problem can be formulated as finding the unique W -band-limited signal f for which $\mathcal{H}_W f = g$. In this direction, we introduce a reproducing kernel \mathcal{Q}_W for which the following equality holds [30]:

$$\forall f \in L^2(\mathbb{R}), \quad \mathcal{P}_W(f) = f, \quad f(t) = \int_{\mathbb{R}} f(s) \mathcal{Q}_W(s, t) ds, \quad t \in [-T, T].$$

Since f is known to be W -band-limited, an explicit form can be derived for \mathcal{Q}_W :

$$\mathcal{Q}_W(s, t) = \int_{-W}^W \exp(2i\pi(t-s)\xi) d\xi = \frac{\sin(2\pi(t-s)W)}{\pi(t-s)} = \mathcal{Q}_W(t, s),$$

which turns out to be the low-pass filtering kernel. The whole procedure can be repeated for this particular choice of \mathcal{Q}_W corresponding indeed to the analytic scaling function being the cardinal sine, ‘‘sinc’’. The eigenfunctions of $\mathcal{H}_W \circ \mathcal{H}_W^*$ are called the Slepian *Prolate Spheroidal wave functions* (PSWF), see for instance [20,33], and are known to realize the maximum energy concentration for a W -band-limited function onto the interval $[-T, T]$. They constitute an orthogonal basis of both $L^2(-T, T)$ and a subspace of $L^2(\mathbb{R})$, known as the Paley-Wiener space of W -band-limited functions. This means in particular that there’s no need for a decomposition like the one of Theorem 3.

This theoretical construction doesn’t lead to a simple numerical algorithm able to generate the W -band-limited extrapolation of g . However, if we denote by \mathcal{P}_{T^c} the orthogonal projector such that $\mathcal{P}_{T^c}(f)(t) = f(t)$ if $|t| > T$, zero in the other cases, we can observe that:

$$\forall t \in \mathbb{R}, \quad g(t) = (\mathcal{P}_T \circ \mathcal{P}_W(f))(t) = (Id - \mathcal{P}_{T^c} \circ \mathcal{P}_W)(f)(t).$$

From [8], we know that the operator $(Id - \mathcal{P}_{T^c} \circ \mathcal{P}_W)$ can be invertible on the space of W -band-limited functions. In this case, $(Id - \mathcal{P}_{T^c} \circ \mathcal{P}_W)^{-1} = \sum_{k \geq 0} (\mathcal{P}_{T^c} \circ \mathcal{P}_W)^k$. An iterative procedure to recover f from g originally due to Gerchberg and Papoulis [11,21] reads:

$$f^{(0)} = g, \quad f^{(1)} = g + \mathcal{P}_{T^c} \circ \mathcal{P}_W(f^{(0)}), \quad \dots, \quad f^{(k+1)} = g + \mathcal{P}_{T^c} \circ \mathcal{P}_W(f^{(k)}). \quad (7)$$

Several proofs of convergence are available, see [31,12,8]. The discrete algorithm is shown to construct the so-called *Minimum-Norm Least-Squares* (MNLS) solution in [6], see also [23].

At this level, it clear how to derive an iterative algorithm for the preceding problem of scale-limited extrapolation following the same stepping stones: it suffices to replace the projector \mathcal{P}_W by \mathcal{P}_{V_j} inside the preceding formula: (see [28])

$$f^{(0)} = g, \quad f^{(1)} = g + \mathcal{P}_{T^c} \circ \mathcal{P}_{V_j}(f^{(0)}), \quad \dots, \quad f^{(k+1)} = g + \mathcal{P}_{T^c} \circ \mathcal{P}_{V_j}(f^{(k)}). \quad (8)$$

This procedure is called the Generalized Gerchberg-Papoulis algorithm (GGP) in [28]. We finally give the main convergence result from [15]:

Theorem 4 (see [15,28]) *Let ϕ be an orthogonal real-valued scaling function and $f^{(k)}$, $k \in \mathbb{N}$ be the sequence of functions generated by the generalized Gerchberg-Papoulis algorithm (8) for some $f \in V_j$, $j \in \mathbb{Z}$. Then, as $k \rightarrow +\infty$, $f^{(k)}$ converges toward the minimum norm solution \tilde{f} defined by:*

- $\tilde{f} \in V_j$
- $\|\tilde{f}\|_{L^2(\mathbb{R})} = \inf \left\{ \|\varphi\|_{L^2(\mathbb{R})}; \varphi \in V_j, \mathcal{P}_T \varphi = \mathcal{P}_T f \right\}$

Recalling Theorem 3, it is likely that the minimum norm solution \tilde{f} belongs to \mathcal{U}_j alone; actually, \tilde{f} is the orthogonal projection of f onto \mathcal{U}_j (see Corollary 1 in [14]).

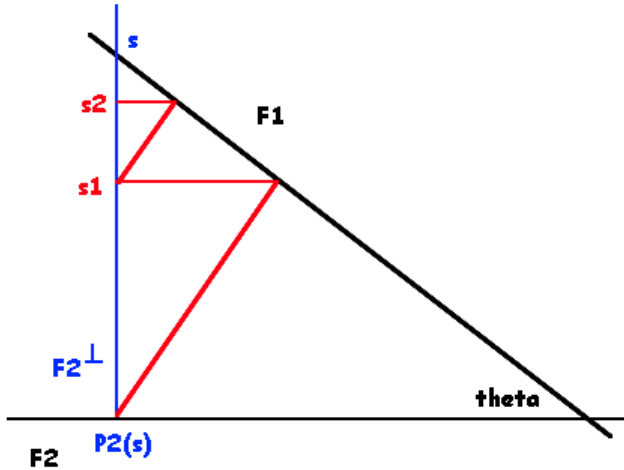


Fig. 2. Illustration of the alternating projections algorithm.

Remark 2 *It is possible to give a direct convergence proof for the band-limited iterative extrapolation algorithm relying on the so-called Alternating Projections Theorem in Hilbert spaces. One defines two subsets of $L^2(\mathbb{R})$: F_1 is the subspace of W -band-limited functions, and F_2 is made of functions vanishing outside of $[-T, T]$. Then, since $F_1 \cap (F_2)^\perp = 0$ as a consequence of the fact that analytic functions have isolated zeros, the iterative procedure (7) converges toward the unique point in F_1 starting from its orthogonal projection on F_2 (see Figure 2). Clearly, the transversality condition isn't met in the general case of scale-limited subspaces (especially because $F_1 \cap (F_2)^\perp = V_j \cap \mathcal{P}_T[L^2(\mathbb{R})]^\perp \neq 0$).*

2.3 Acceleration by Conjugate Gradient resolution

The iterative procedures from the preceding section can be readily implemented on a computer; however, it is a well-known fact that they are endowed with rather poor practical convergence properties. In order to fix this drawback, several acceleration routines have been proposed: we shall concentrate onto the ones based on the Conjugate Gradient algorithm (see [12,26] for the band-limited extrapolation and [15] for the scale-limited one). The idea is the same in both cases of band-limited and scale-limited extrapolation: one considers the pointwise collection of sampled values for $g(t)$, and while seeking for an approximation of $f(t)$, aim at solving efficiently the linear system given by $g(t) = \mathcal{P}_T \circ \mathcal{P}_{V_j}(f)(t)$ (for band-limited extrapolation, one inserts \mathcal{P}_W in place of \mathcal{P}_{V_j} in the former formula).

In this direction, it is necessary to define $\vec{g} \in \mathbb{R}^M$ as the collection of N observed values (the samples of $g(t)$) completed by zero ($M > N$), the discrete operators T_N and P_j being respectively the diagonal truncation matrix and the discrete projector onto V_j ; the unknown is therefore $\vec{f} \in \mathbb{R}^M$ such that $\vec{g} = T_N P_j \vec{f}$. Following [15], we begin by observing that these matrices have several interesting properties: in case the base of V_j given by the scaling function ϕ is orthonormal, the matrix P_j is symmetric. Moreover, we clearly have that $(T_N P_j)^* = P_j T_N$ and $T_N^* T_N = T_N$. The Conjugate Gradient method is to be applied to the *normal equations* yielding the least-squares solution of the discrete extrapolation problem:

$$\bar{x} = \inf_{x \in \mathbb{R}^M} \|Ax - b\|_{L^2} \Leftrightarrow (A^* A)\bar{x} = A^* b.$$

In the present case, this reduces to find efficiently \vec{f} solving

$$\begin{aligned} (T_N P_j)^* \vec{g} &= P_j T_N \vec{g} = P_j g = (T_N P_j)^* (T_N P_j) \vec{f} \\ &= P_j^* T_N^* T_N P_j \vec{f} \\ &= P_j T_N P_j \vec{f} \\ &= P_j T_N \vec{f} \in \mathcal{U}_j, \end{aligned}$$

see Theorem 3. Intermediate steps are done by noting that $T_N \vec{g} = \vec{g}$ and $P_j \vec{f} = \vec{f}$ by their very definition. Since $P_j T_N P_j$ is symmetric positive definite (observe that $x^T P_j T_N P_j x = (P_j x)^T T_N (P_j x) > 0$ for a smooth scaling function ϕ in case (6) holds), the Conjugate Gradient algorithm can be set up to solve in a finite number of operations the linear system: $b := P_j g = P_j T_N P_j \vec{f} := Ax$. We recall it briefly:

Fix $r_0 = b$ and $p_0 = r_0$; while $r_k > \varepsilon$,

- $q_k = Ap_k$
- If $k=1$ $p_1 = r_0$, else

$$\beta_{k-1} = \frac{r_{k-1}^\top r_{k-1}}{r_{k-2}^\top r_{k-2}}$$

$$p_k = r_{k-1} + \beta_{k-1} p_{k-1}$$
- $\alpha_k = \frac{r_{k-1}^\top r_{k-1}}{p_k^\top q_k}$
- $x_k = x_{k-1} + \alpha_k p_k$
- $r_k = r_{k-1} - \alpha_k q_k$

This algorithm will hopefully converge in a finite number of iterations toward the vector $x \simeq \vec{f}$ which contains the sampled values of the minimum-norm solution \vec{f} of the scale-limited extrapolation problem up to a given (small) tolerance $0 < \varepsilon \ll 1$. We close this section by noticing that in practice, it is not necessary to implement the matrix P_j : the WaveLab package contains commands for computing the Fast Wavelet Transform and its inverse. It suffices thus to use it, to cut the coefficients corresponding to scales finer than j , and to apply the inverse transform. For big values of M , it will also result in a quicker execution.

3 Application to stock market indexes

One theoretical obstruction for considering stock market indexes in the context of multi-resolution analysis of $L^2(\mathbb{R})$ is that these signals aren't typically functions of finite energy defined on the whole real line. Hence a way around this issue is to extend them with zero (*zero-padding*) outside of their definition interval (which is obviously compact); the resulting function belongs to $L^2(\mathbb{R})$ and we can apply the preceding developments. Moreover,

if considering time scales not shorter than daily quotations, stock market signals can be treated as Lipschitz continuous $W^{1,\infty}(\mathbb{R})$ functions (with possibly big Lipschitz constants depending on volatility). These are the only hypotheses really necessary to be made on the signal in order to set up the preceding algorithms.

3.1 The need for wavelet bases on a bounded interval

As a consequence of the aforementioned minor issue concerning our signals, it may result difficult to set up the usual wavelet algorithms; in particular, since stock indexes are strictly positive and strongly non-periodic, extending them either by zero or periodically will result in big discontinuities on the border of the interval of observations. This leads typically to the emergence of big and spurious wavelet coefficients at the high resolution scales which render for these discontinuities (in practice, standard algorithms for wavelet transforms² extend the source signal to the whole real line periodically). Hence, in order to be able to get a correct compression of such a signal by projecting it onto a scale-limited space, the use of *wavelet bases on a bounded interval* became absolutely necessary.

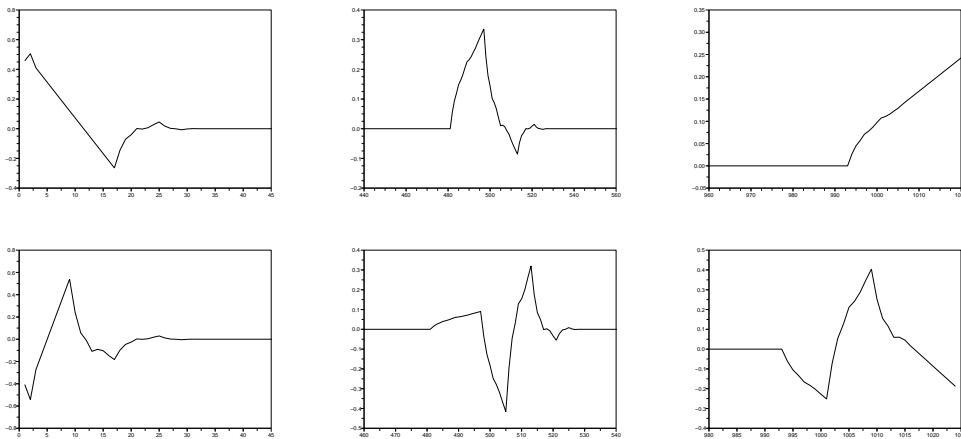


Fig. 3. Scaling function (top) and its mother wavelet (bottom) in the interval.

The theory of multi-resolution on a bounded interval has been developed after the pioneering work by Meyer [16]. However, his construction couldn't be used to produce a robust numerical algorithm. This problem has been fixed later by Cohen, Daubechies, Jawerth and Vial [3,5]; see also *e.g.* [4,17,19,13]. This results in a local modification of the scaling functions at the edge of the interval so that all of the properties enjoyed by the wavelets on \mathbb{R} are preserved. In particular, as used in the characterization of spaces $C^s(\mathbb{R})$ via the decay of wavelet coefficients, it is important to preserve both the wavelet's smoothness and its vanishing moments. This corresponds to introducing modified filters (the sequences usually denoted by $\{h_n\}$ and $\{g_n\}$ in the cited papers) for the treatment of the edges of the interval; otherwise, away from the borders, the algorithm remains essentially the same as for the case of the infinite real line. On Figure 3, one can visualize in the middle column the same scaling function and mother wavelet than in Figure 1; however, in both left and right columns, their modifications able to handle properly discontinuities possibly appearing at the edges of a finite interval are displayed.

² This work has been done relying on the free package WaveLab developed for MatLab by Dave Donoho and co-workers, see <http://www-stat.stanford.edu/wavelab/>. However, we used its conversion for the free software SciLab, see <http://www.scilab.org>.

We stress the fact that, since we aim at extrapolating data on the right edge of the observations interval by projecting the whole set onto a scale-limited subspace, it is of crucial importance to prevent the apparition of high wavelet coefficients in this region because of artificial periodicity. In this last case, the scaling function coefficients kept when annihilating all the others corresponding to finer scales are likely to be modified because of this “Gibbs-like” phenomenon. Consequently, the initial data given to the GGP algorithm (8) would be rather inaccurate, and would thus result in a poor extrapolation since such an inverse problem is quite sensitive to noise (see [14]).

3.2 Multi-scale decomposition and scale-limited extrapolation

Hereafter, the notation $I(t) \in \mathbb{R}^+$ will refer to some nonnegative stock index and \vec{I} , to its sampled values on a time-discrete grid. In practice, we shall always rely on daily quotations, which are freely available at <http://fr.finance.yahoo.com>. The general framework presented in this paper can be set up by first, selecting a multi-resolution analysis on a bounded interval which is compatible with I (in the sense that base functions should have vanishing moments and smoothness properties compatible with the volatility of the index under consideration), then choosing a scale $j \in \mathbb{Z}$ which realizes a good balance between accuracy and data compression. With both these elements in hand, it is possible to perform short-time extrapolation as explained in the previous sections on real-life examples.

According to the WaveLab package, only the standard compact support Daubechies wavelet with either 2 or 3 vanishing moments is available (see again Figure 3 for an illustration). For most of the indexes, the one with only 2 vanishing moments gave the best results for observations ranging from 512 to 2048 days of quotations (recall that one year gives roughly 250 quotation days). Once the multi-resolution analysis has been selected, the remaining issue is to choose a scale parameter $j \in \mathbb{Z}$ which retains enough information in the time series, but is low enough to allow for a non-trivial extrapolation. This means that $\|\mathcal{P}_{V_j} I - I\|_{L^2}$ should be small; in other words, the wavelet coefficients associated to subspaces W_ℓ with $\ell > j$ should be rather small (the smaller, the better). In this case, one can say that the index I is well compressed in the base $\phi_{j,n}$ of V_j . This has to do with the fact that oscillations in the index at a scale finer than j don't really carry important information; volatility (in the sense formerly mentioned of standard deviation) may exist, but what matters is that its effect onto the global trend of I is very limited in such a situation.

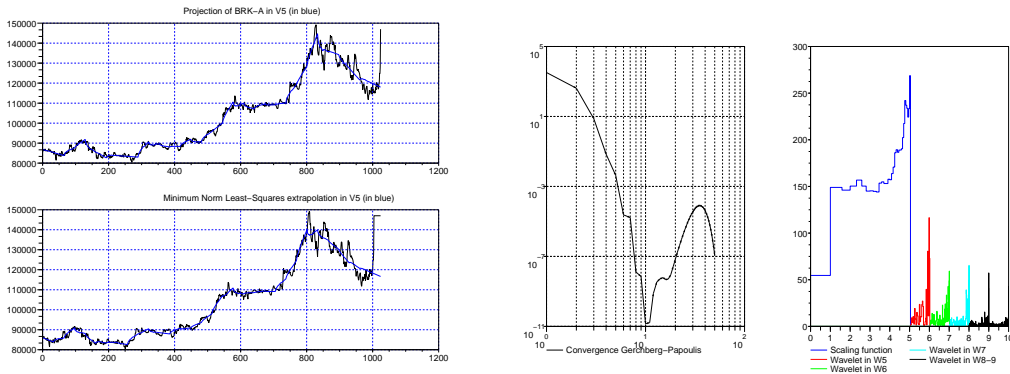


Fig. 4. Quotations of Berkshire-Hathaway A (left) and wavelet coefficients (right).

This favorable context doesn't happen systematically: on Figure 4, we display the 1024-days evolution (ending on october 17, 2008), wavelet coefficients and conjugate gradient convergence which produce a 20-days extrapolation for the famous Buffett's holding, Berkshire-Hathaway. In this rather spectacular case, one sees clearly that during most of the observation interval, prices follow rather well the trend obtained by projecting them onto V_5 ;

however, in the last days, a very rapid growth cancels the validity of such an approximation in a scale-limited space. The consequences are easily seen on the right figure: in blue are represented the coefficients of $\phi_{j=5,n}$. In red, green, light blue and black are represented the wavelet coefficients in W_5 , W_6 , W_7 and W_8 , W_9 multiplied by 10 respectively (the abscissa is the scale index). Obviously, for this pathological example, we clearly have important energy concentrated in the fine scales: hence the conjugate gradient algorithm converges poorly, and the resulting extrapolation is quite poor because the scale-limited space doesn't really match all the meaningful features of the evolution of the share price. As a consequence, working on indexes is safer than to do so on individual share's prices, mainly because the former are more likely to display a smoother behavior because they are computed as convex combinations of the latter.

4 Numerical experiments on real data

4.1 A quiet bull market

We begin by considering the French leading index CAC 40 on a period of 1024 days (roughly 4 years beginning in 2003) covering the end of the bear market which resulted from the "internet krach" and the September 11 events. The very low interest rates, cut by FED Chairman Alan Greenspan, helped to sustain the prices and allowed for a bull market to take place. It was characterized by a moderate growth and a low volatility; this is easily noticeable by looking at the wavelet coefficients in the middle/right part of Figure 5 since the coefficients in W_j , $j > 5$ look quite small even if they are again multiplied by 10.

This is the best context in which to set up our scale-limited extrapolation algorithms. For a 30 days forecast (roughly one month and a half), one gets a monotone convergence of the conjugate gradient routine in less than 10 iterations. On the top graphic of Figure 5, one sees that starting from the red arrow (at the abscissa $t = 1004$), the MNLS extrapolated part matches quite well the evolutions of the index. Obviously, one cannot hope to obtain *exactly* the real oscillations, as what is sought is an approximation in V_5 only, but the idea is more to extrapolate a general trend and to define a *confidence interval* around it which should contain the fluctuations of the near-future.

In order to set up this confidence interval, one considers the discarded part of the signal (see bottom left of Figure 5), namely $I(t) - \mathcal{P}_{V_j} I(t)$ (which has some vanishing moments because it belongs to the wavelet spaces $\cup_k W_k$, $k \geq j$), computes its distribution function together with its standard deviation. In case it would be a standard Gaussian distribution (see for instance the bottom of Figure 5), then it would suffice to trace both red curves at a distance of 2 times the standard deviation to ensure a 95% confidence interval. In realistic experiments, we found that the discarded signal approaches a Gaussian distribution when overall volatility was low thus one can only speak about an *approximated confidence interval*. The stopping criterion ε for the CG algorithm was 10^{-9} .

4.2 A moderately volatile bear market

During years 2003-2007, not only the shares prices have tremendously increased; real-estate market value has inflated more than stocks in certain areas as a consequence of wide and easy credit allocation. Some of the gains obtained on financial markets have also been recycled in "concrete" during this period. On Figure 6, we are interested in the FTSE EPRA Eurozone index (EPRA meaning European Real Estate Association³), which contains es-

³ See <http://www.epra.com>

essentially shares linked to european real-estate markets, like banks, insurance companies and promoters. Obviously, starting from august 2007, this index entered into a bear market as a consequence of the so-called *subprime crisis*⁴ ignited in the USA in february 2007 with the depreciations of the bank HSBC.

We work with a measures interval of 512 days (roughly 2 years, ending in february 2008) mainly because this index is too recent to allow for a 1024 days sample. A bear market being in general more volatile than a bull one (one says that *shares prices decrease twice as fast as they grow* [6]), the Gaussian approximation of the “noise” defined as $I(t) - \mathcal{P}_{V_j} I(t)$ with $j = 4$ is less convincing; however, the approximated confidence interval remains rather satisfying even in the clearly decreasing period ranging from day 200 to 500. From the repartition of the energy in the wavelet coefficients (middle in Figure 6, it may sound reasonable to work in the V_5 space rather than in V_4 as we did. However, this wouldn’t lead to a sensible improvement in the quality of the results. On the top of Figure 6, one sees that the MNLS 15 days extrapolation (starting from the red arrow at the abscissa $t = 497$) is rather good in determining a trend around which a reliable confidence interval can be built. The convergence of the CG algorithm was satisfying; the stopping criterion ε for the CG algorithm was 10^{-10} .

4.3 What happens for a krach ?

On Figure 7, we display what happened on the Chinese index Hang-Seng in the first days of october 2008: namely, because of recession fears in both the USA and the Eurozone, operators sold massively emerging countries shares regardless of the so-called *decoupling theory*⁵. The index under consideration lost around 25% of its value in less than 10 days and authorities decided to close the market several times and to ban short-selling in order to prevent a complete collapse. In such a case, there’s no hope that the scale-limited extrapolation (starting from the red arrow at the abscissa $t = 1004$) can give a reliable result (see top of Figure 7); however, what can be considering as interesting is that such an analysis highlights the fact that the discarded part of the signal (see bottom of Figure 7) has a now repartition function which is different from the standard Gaussian: it is quite peaked which a thick tail on the left side. Hence this suggests a potentially volatile index which is likely to have violent corrections. Increasing the value of j may improve the result by including more fine scale features inside $\mathcal{P}_{V_j} I(t)$: however, a strong reduction of the size of the extrapolation interval would also come as a direct consequence of (6).

5 Conclusion and outlook

We presented in this paper a quite general methodology to handle a standard problem in Empirical Finance, that is, trying to extrapolate stock market indexes during periods of mild volatility relying on the nowadays well-known theory of multi-resolution and discrete wavelet subspaces of $L^2(\mathbb{R})$. The numerical results obtained on real-life data coming from daily quotations indicate that this process can be considered as reliable as long as the basic assumption of the relevance of a scale-limited approximation $\mathcal{P}_{V_j} I$ of the index I holds. Especially, during strong corrections (like *e.g.* in october 2008), one notices rather easily that $I - \mathcal{P}_{V_j} I$ ceases to be small and is even not distributed like a standard Gaussian density of the same average and variance. When this “empirical density” is very peaked (like in Figure 7), the scale-limited extrapolation cannot be taken as a meaningful possible forecast for the index. Concerning the size of the extrapolation interval, it depends on the level of resolution j of the scale-limited approximation (a lower j gives a longer extrapolation, but

⁴ See http://en.wikipedia.org/wiki/Subprime_mortgage_crisis

⁵ See http://en.wikipedia.org/wiki/Decoupling#Decoupling_and_the_stock_market_declines_of_January_2008

at the price of a poorer accuracy) and on the smoothness of the scaling function (smoother yields a bigger interval). The extreme case being the one of analytic functions from which one can theoretically extrapolate onto the whole real line \mathbb{R}).

A Treating the very oscillating “log-returns” function

There exists another way, maybe more natural, to extrapolate a stock index $I(t)$: it is to work with its so-called “log-returns” function (see *e.g.* [6]) defined as follows:

$$S(t) = \ln \left(\frac{I(t)}{I(t-1)} \right) = \frac{\ln(I(t)) - \ln(I(t-1))}{1} \simeq \frac{d}{dt} \ln(I(t)) = \frac{I'(t)}{I(t)}. \quad (\text{A.1})$$

As the quotations are taken on a daily base, a straightforward piecewise linear interpolation produces an overall positive Lipschitz continuous function belonging to the Sobolev space $W^{1,\infty}$. Therefore, extracting its logarithmic derivative makes sense at least within the framework of L^∞ , or maybe even in functions of bounded variation. The advantage is that, since S is a very oscillating function around zero, the need for multi-resolution bases defined on a bounded interval may be relaxed as (in general) one shouldn’t expect too big variations in the L^∞ norm. Hence the discontinuities possibly showing up at the edges of the interval of observations can be really smaller compared to the ones arising when working with the index I itself. This can open the way to using much smoother wavelet bases, like for instance the Symmlet: see Figure A.1. However, a small issue arises because, since all the filtering and the extrapolation is done on the logarithmic derivative (A.1), the starting point $I(0)$ is lost; indeed, by definition, there holds:

$$\forall t \geq 0, \quad I(t) = I(0) \exp \left(\int_0^t S(s).dx \right).$$

After the separation of S between low and fine scales in order to produce $\mathcal{P}_{V_j}(S)$, $I(0)$ becomes useless, but one can use another value. Let C_j be the support of the scaling function $\phi_{j,n=0}(t)$, as an initial point, we take:

$$\tilde{I}_0 = \frac{1}{|C_j|} \int_{C_j} I(s).ds.$$

This allows to compute a scale-limited index at level j like:

$$\forall t \geq 0, \quad \tilde{I}_j(t) := \tilde{I}_0 \exp \left(\int_0^t \mathcal{P}_{V_j}(S)(s).dx \right).$$

We applied these ideas to the example of EPRA Eurozone (see §4.2); on the top of Figure A.2, we display the multi-resolution decomposition of its log-returns $S(t)$ ranging from V_3 , W_3 up to W_5 . On the middle, the extrapolation of the index obtained through the scale-limited extrapolation of its log-returns function is shown (starting from the red arrow at abscissa $t = 497$). It is considerably smoother than the result from Figure 6; the first reason is that the analyzing wavelet is smoother, the second one is that moreover, S is integrated to deduce I . On the bottom, one can see that the convergence of the Conjugate Gradient algorithm is satisfying even for this type of extrapolation problem. However, when looking at the wavelet coefficients, one may think it became more difficult to justify why the space V_4 can be chosen to approximate the function $S(t)$ as the fine scale coefficients aren’t small like in the examples shown in the preceding sections. But since we work now with a derivative, big values on a fine scale correspond to both important and rapid (maybe erratic) variations of the index $I(t)$. It is therefore perfectly normal to remove them in order to derive a smooth trend; the working subspace V_j should correspond now to a set of moderately big coefficients in the discrete wavelet transform of S .

References

- [1] M. Aminghafari, J.M. Poggi, *Forecasting time-series using wavelets*, Int. J. Wavelets, Multiresolution and Information Proc. **5** (2007) 709–724.
- [2] Miguel A. Ario, *Forecasting Time Series via Discrete Wavelet Transform*, Computing in Economics and Finance (1996) 005, Society for Computational Economics (<http://www.unige.ch/ce/ce96/welcome.html>).
- [3] A. Cohen, I. Daubechies, B. Jawerth, P. Vial, *Multiresolution analysis, wavelets and fast algorithms on an interval*, C. R. Acad. Sci. Paris **316** Série I (1993) 417–421.
- [4] J. Cheng, W.-C. Shann, *About wavelets on the interval*, in “Proceedings of the International Mathematics Conference '94,” World Scientific (1995) 49–64.
- [5] A. Cohen, I. Daubechies, P. Vial, *Wavelets on the interval and fast wavelet transforms*, Applied and Computational Harmonic Analysis **1** (1993) 54–81.
- [6] R. Cont, *Empirical properties of asset returns: stylized facts and statistical issues*, Quantitative Finance **1** (2001) 223–236.
- [7] M.J. Croft, J.A. Hogan, *Wavelet-based signal extrapolation*, Fourth International Symposium on Signal Processing and Its Applications, ISSPA 96, **2** (25-30 Aug 1996) 752-755.
- [8] D.L. Donoho, P.B. Stark, *Uncertainty principles and signal recovery*, SIAM J. Appl. Math. **49** (1989) 906–931.
- [9] P. Fryzlewicz, S. Van Bellegem, R. Von Sachs, *Forecasting non-stationary time series by wavelet process modeling*, Annals of the Institute of Statistical Mathematics **55** (2003) 737–764.
- [10] R. Gençay, F. Selcuk, B. Whitcher, **An Introduction to Wavelets and Other Filtering Methods in Finance and Economics** (2002) Academic Press, New York.
- [11] R.W. Gerchberg, Super-resolution through error energy reduction. Opt. Acta **21**, 709-720 (1974)
- [12] A.K. Jain, S. Ranganath, *Extrapolation algorithm for discrete signal with application in spectral estimation*, IEEE Trans. Acoust. Speech Signal Process. **29** (1981) 830–845.
- [13] Wei Siong Lee, Ashraf A. Kassim, *Signal and Image Approximation using Interval Wavelet Transform*, IEEE Trans. Imag. Process. **16** (2007) 46–56.
- [14] L.-C. Lin, C.-C. Jay Kuo, *On theory and regularization of scale-limited extrapolation*, Signal Processing **54** (1996) 225–237.
- [15] L.-C. Lin, X.-G. Xia, C.-C. Jay Kuo, *On the convergence of wavelet-based iterative signal extrapolation algorithms*, Signal Processing **48** (1996) 51–65.

- [16] Y. Meyer, *Ondelettes sur l'intervalle*, Riv. Matem. Iberoamericana **7** (1991) 115–133.
- [17] S. Mallat, **A wavelet tour of signal processing**, Academic Press, 1998.
- [18] R. J. Marks II, *Alternating Projections onto Convex Sets*, in “Deconvolution of Images and Spectra” (Second Edition), Peter A. Jansson Editor, 1997, 1984 by Academic Press
- [19] K. Melkemi, Z. Mokhtari, *Wavelet bases on the interval: a new approach*, Int. J. Math. Anal. **1** (2007) 1019–1030.
- [20] I.C. Moore, M. Cada, *Prolate Spheroidal wave functions, an introduction to the Slepian series and its properties*, Appl. Comput. Harmonic Anal. **16** (2004) 208–230.
- [21] A. Papoulis, *A New Algorithm in Spectral Analysis and Band-Limited Extrapolation*, IEEE Trans. Cir. & Sys., **22** (1975) 735–742.
- [22] D.B. Percival, A.T. Walden, **Wavelet Methods for Time Series Analysis**, (2000) Cambridge University Press.
- [23] L.C. Potter, K.S. Arun, *Energy concentration in band-limited extrapolation*, IEEE Trans. Acoustics, Speech and Signal Processing **37** (1989) 1027–1041.
- [24] J.B. Ramsey, *The contribution of wavelets to the analysis of economic and financial data*, Phil. Trans. R. Soc. Lond. A **357** (1999) 2593–2606.
- [25] O. Renaud, J.-L. Stark, F. Murtagh, *Wavelet-based Forecasting of Short and Long Memory Time Series*, Signal Processing **48** (1996) 51–65.
- [26] T. Strohmer, *On discrete band-limited signal extrapolation*, Contemp. Math. **190**:323–337, 1995
- [27] M. Unser, *Sampling—50 years after Shannon*, Proc. IEEE **88** (2000) 569–587.
- [28] X.-G. Xia, C.-C. Jay Kuo, Z. Zhang, *Signal extrapolation in wavelet subspaces*, SIAM J. Sci. Comp. **16** (1995) 50–73.
- [29] X.-G. Xia, L.-C. Lin, C.-C. Jay Kuo, *On the limit of sampled signal extrapolation with a wavelet signal model*, Numer. Funct. Anal & Optimiz. **16** (1995) 1255–1272.
- [30] K. Yao, *Applications of reproducing kernel Hilbert spaces—Band-limited signals models*, Information & Control **11** (1967) 429–444.
- [31] D.C. Youla, H. Webb, *Image Restoration by the Method of Convex Projections: Part 1 Theory*, IEEE Trans. Med. Im., vol. MI-1, No. 2, 1982.
- [32] Peng Yu, Anna Goldenberg, Zhiqiang Bi, *Time Series Forecasting using Wavelets with Predictor-Corrector Boundary Treatment*, in 7th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (2001). (<http://www.cs.cmu.edu/~anya/papers/dbfinal.ps>)
- [33] A. Zayed, *A generalization of the prolate spheroidal wave functions*, Proc. Amer. Math. Soc. **135** (2007) 2193–2203.

[34] E. Zeidler, **Applied Functional Analysis**, Applied Mathematical Sciences
108, Springer Verlag 1995.

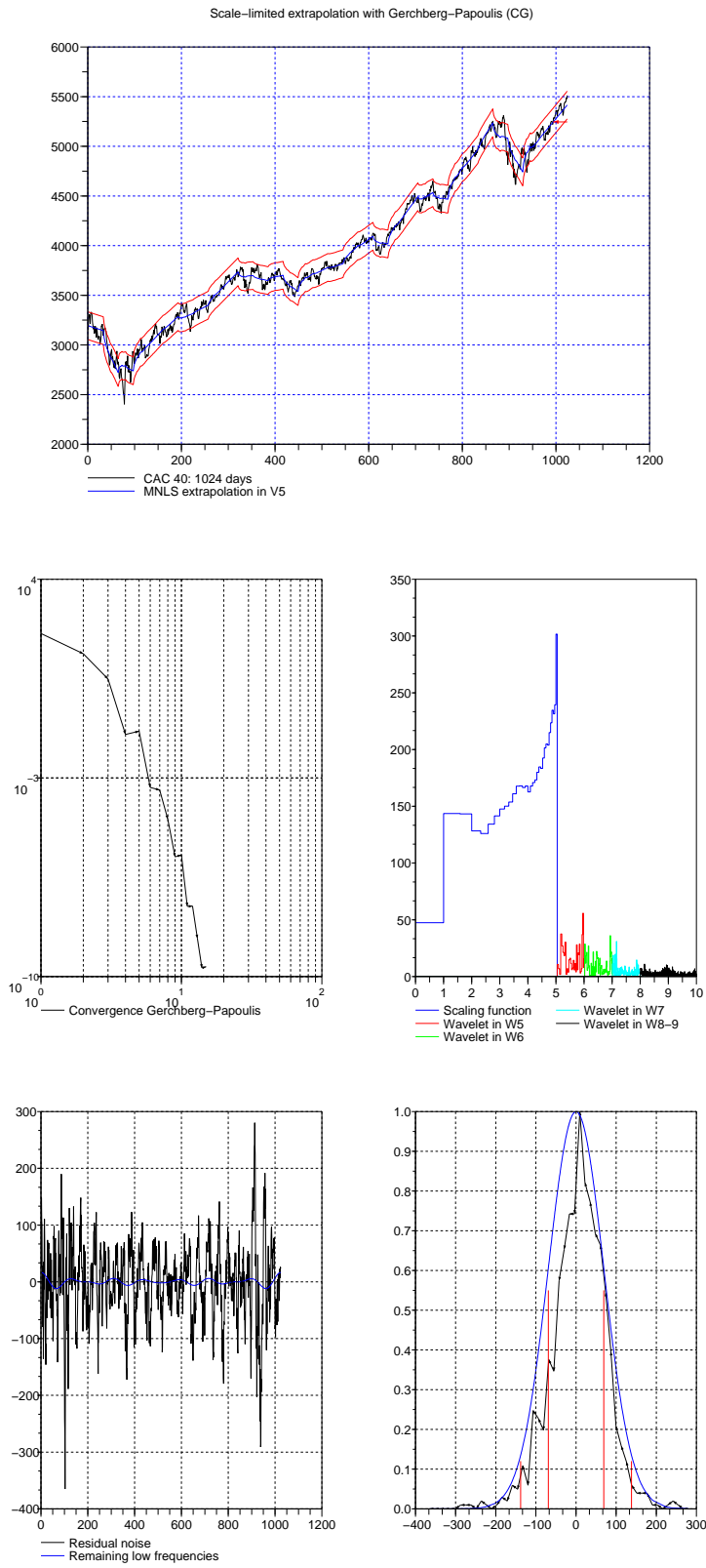


Fig. 5. CAC 40 index on 1024 days: 30 days extrapolation (top), scale-limited decomposition and convergence (middle) and remaining noise repartition (bottom).

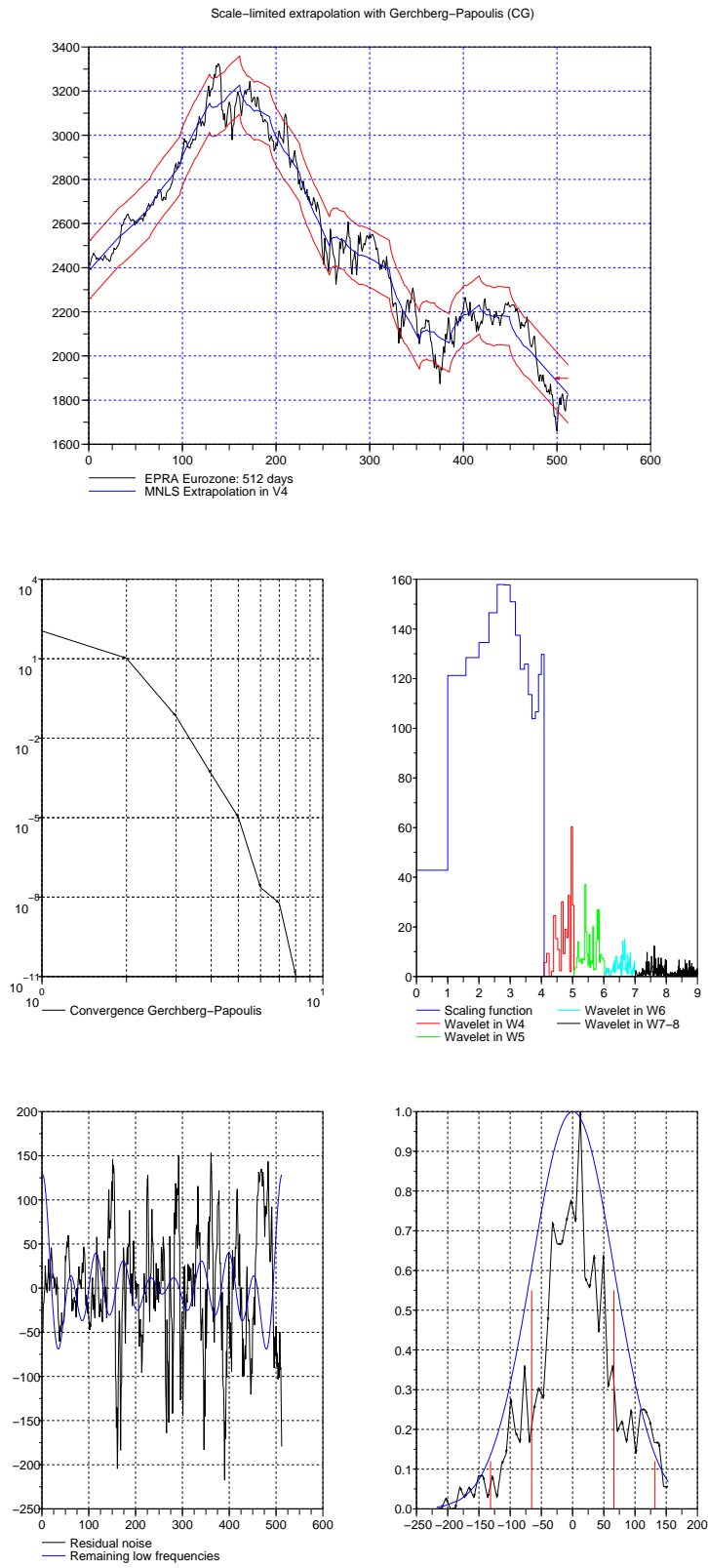


Fig. 6. EPRA index on 512 days: 15 days extrapolation (top), scale-limited decomposition and convergence (middle) and remaining noise repartition (bottom).

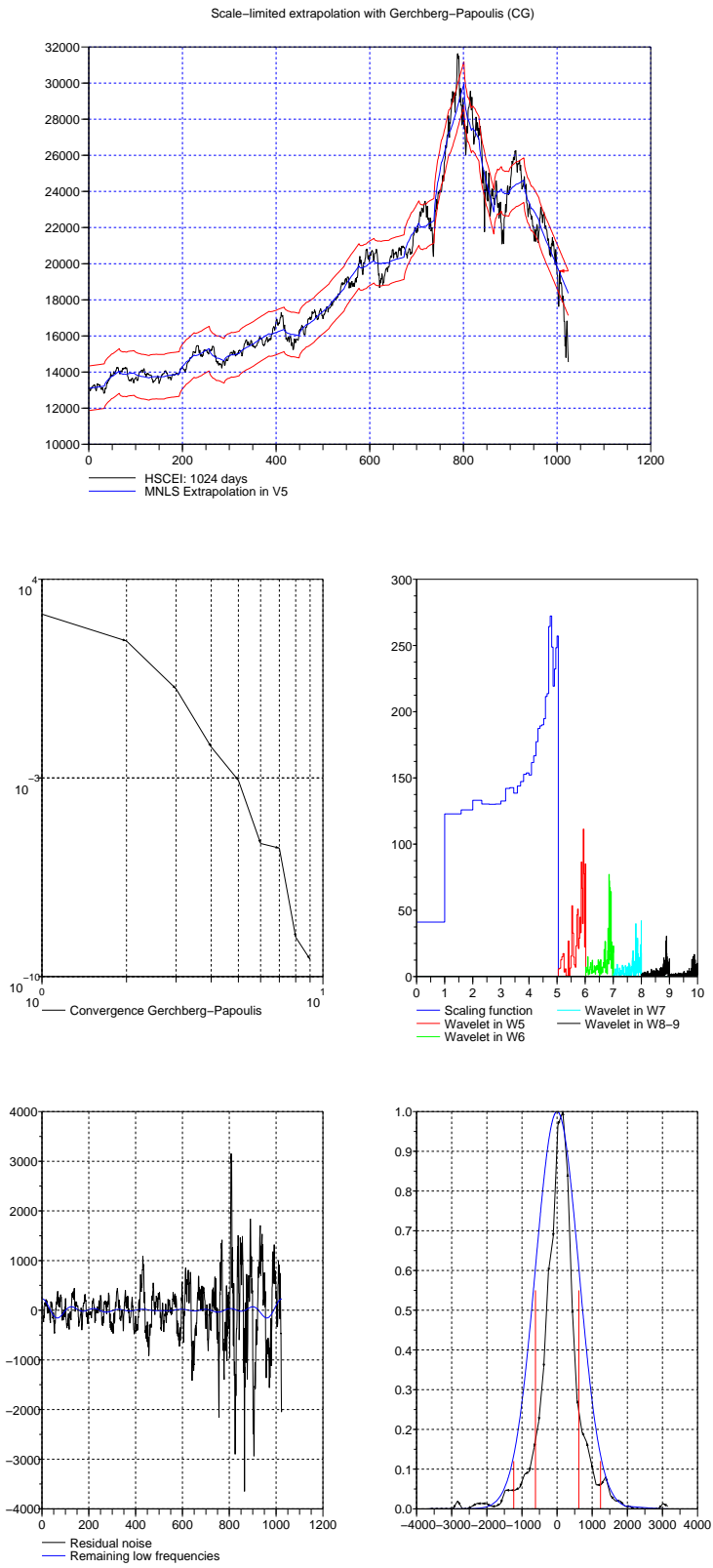


Fig. 7. HSCEI index on 1024 days: 20 days extrapolation (top), scale-limited decomposition and convergence (middle) and remaining noise repartition (bottom).

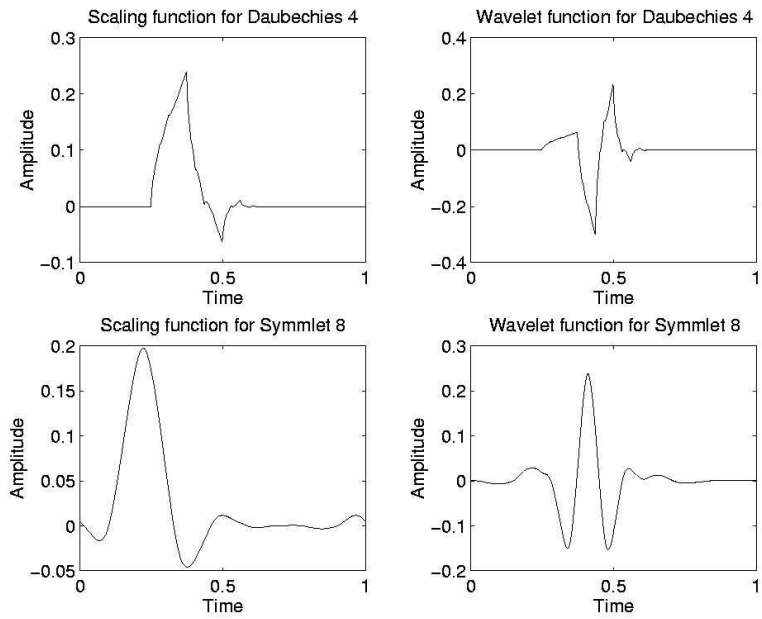


Fig. A.1. Comparison between the Daubechies 4 (top) and Symmlet 8 (bottom) basis functions.

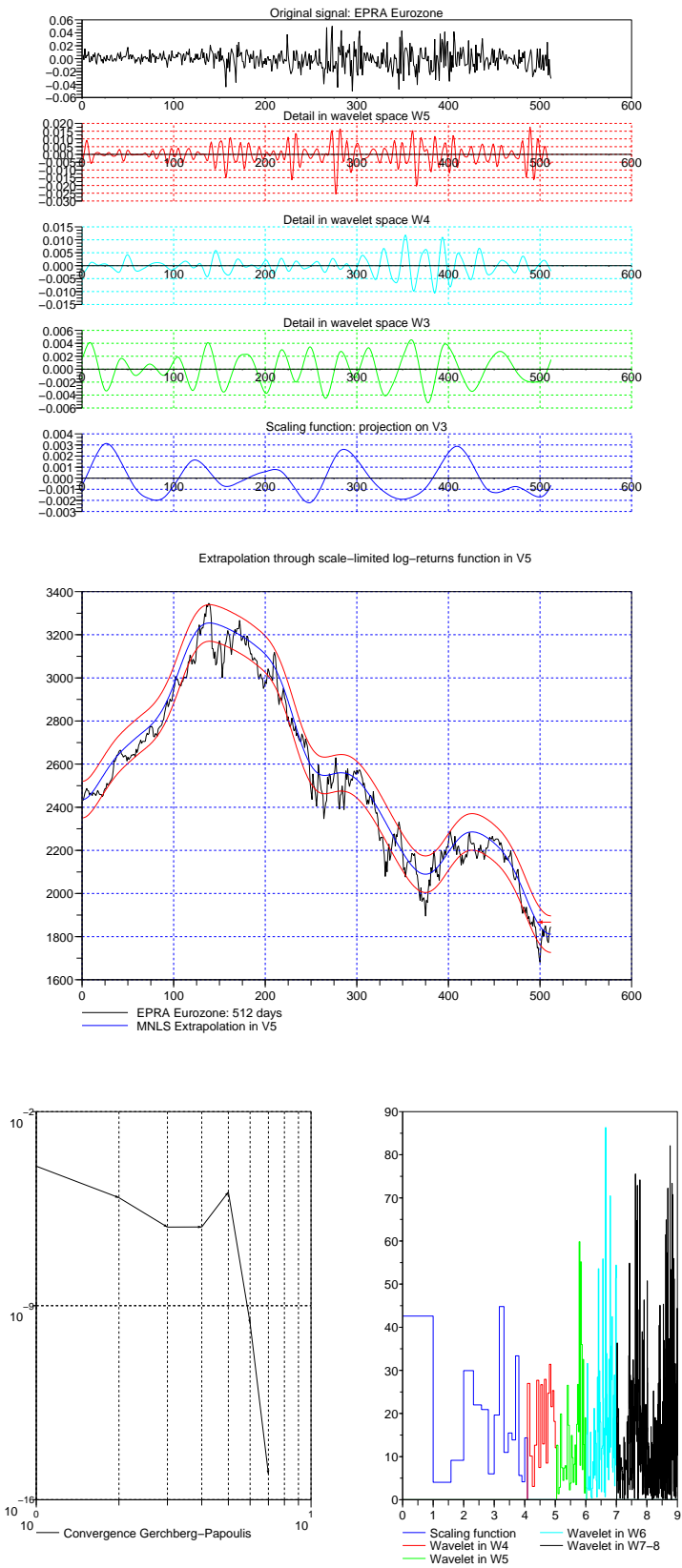


Fig. A.2. EPRA on 512 days: 15 days extrapolation (top), scale-limited decomposition and convergence (middle) and remaining noise repartition (bottom).