The Continuous Wavelet Transform: moving beyond uni and bivariate analysis* †

Luís Aguiar-Conraria‡ Maria Joana Soares§

October 13, 2011

Abstract

Economists are already familiar with the Discrete Wavelet Transform. However, a body of work using the Continuous Wavelet Transform has also been growing. We provide a self-contained summary on continuous wavelet tools, such as the Continuous Wavelet Transform, the Cross-Wavelet, the Wavelet Coherency and the Phase-Difference. Furthermore, we generalize the concept of simple coherency to Partial Wavelet Coherency and Multiple Wavelet Coherency, akin to partial and multiple correlations, allowing the researcher to move beyond bivariate analysis. Finally, we describe the Generalized Morse Wavelets, a class of analytic wavelets recently proposed. A user-friendly toolbox, with examples, is attached to this paper.

Keywords: Continuous Wavelet Transform, Cross-Wavelet Transform, Wavelet Coherency, Partial Wavelet Coherency, Multiple Wavelet Coherency, Wavelet Phase-Difference; Economic fluctuations.

1 Introduction

Economic agents simultaneously operate at different horizons. For example, central banks have different objectives in the short and long run, and operate simultaneously at different frequencies (see Ramsey and Lampart 1998a). More than that, many economic processes are the result of the actions of several agents, who have different term objectives. Therefore, economic time series are an aggregation of components operating on different frequencies. Several questions about the data are connected to the understanding of the time series behavior at different frequencies.

Fourier analysis allows us to study the cyclical nature of a time series in the frequency domain. In spite of its utility, however, under the Fourier transform, the time information of a time series is lost. Because of this loss of information, it is hard to distinguish transient relations or to identify

*Luís Aguiar-Conraria acknowledges financial support from Fundação para a Ciência e a Tecnologia, project "Oil shocks and the Macroeconomy: Econometric estimation, economic modeling and policy implications", PTDC/ECO/64750/2006. We thank three anonymous referees for insightful comments and criticisms as well as the editor for helpful suggestions.

†There is a Matlab toolbox associated with this paper, called ASToolbox, which is available at http://sites.google.com/site/aguiarconraria/joanasoares-wavelets.

‡NIPE and Economics Department, University of Minho. E-mail: lfaguiar@eeg.uminho.pt

§NIPE and Department of Mathematics and Applications, University of Minho. E-mail: jsoares@math.uminho.pt

when structural changes do occur. Moreover, these techniques are only appropriate for time series with stable statistical properties, i.e. stationary time series.

As an alternative, wavelet analysis has been proposed. Wavelet analysis performs the estimation of the spectral characteristics of a time series as a function of time, revealing how the different periodic components of the time series change over time. As we will see, one major advantage afforded by the wavelet transform is the ability to perform natural local analysis of a time series: the wavelet stretches into a long function to measure the low frequency movements, and it compresses into a short function to measure the high frequency movements.

The pioneering work of Ramsey and Lampart (1998a and 1998b) and Ramsey (1999 and 2002) was followed by Gençay, Selçuk and B. Withcher (2001a, 2001b and 2005), Wong, Ip, Xie and Lui (2003), Connor and Rossiter (2005), Fernandez (2005) and Gallegati and Gallegati (2007), among others. All these works, however, have one common characteristic. They all rely on the discrete wavelet transform (DWT).[1]

More recently, tools associated with the continuous wavelet transform (CWT) are becoming more widely used. Raihan, Wen and Zeng (2005), Jagrič and Ovin (2004), Crowley and Mayes (2008), Aguiar-Conraria, Azevedo and Soares (2008), Baubeau and Cazelles (2009), Rua and Nunes (2009), Rua (2010) and Aguiar-Conraria and Soares (2011a and b); provide some examples of useful economic applications of these tools. In his review, Crowley (2007) also includes an introduction to the continuous wavelet transform.[2]

As it will become clear, the continuous wavelet transform maps the original time series, which is a function of just one variable — time — into a function of two variables — time and frequency, providing highly redundant information. This suggests that it might be possible to compute the wavelet transform for just a "careful" selection of values of the frequency and time parameters and still not loose any information (by this we mean that it is possible to recover the original time series from its transform). This is basically the idea of the DWT: we just compute the transform for a very special discrete choice of the parameter values for time and frequency. This leads to a very simple and efficient iterative scheme to compute the transform. The simplicity, ease of implementation and, most of all, its very low computational effort justifies the popularity of the DWT. Also, the DWT is very useful when our main concern is to compress information, as it is very often the case in signal or image processing. The most common transform among economists is probably the so-called MODWT (maximum overlap discrete wavelet transform). The MODWT can be seen as a kind of compromise between the DWT and the CWT; it is a redundant transform, because while it is efficient with the frequency parameters it is not selective with the time parameters, but not as redundant as the CWT. The redundancy of the CWT has a price — essentially, the price is paid in computational time — but it has also some advantages. One of the advantages lies on the fact that it gives us a large freedom in selecting our wavelets, whilst this choice is more limited in the discrete setting. But, most of all, in our opinion, the redundancy makes it much easier to interpret the results obtained and to draw conclusions from the data. The pictures obtained with the CWT are much easier to interpret than the results obtained with the DWT. Hence, if we are not concerned

[1] For an excellent review on discrete wavelet applications in economics, see Crowley (2007).

[2] Crowley (2010) provides an account not only on discrete and continuous wavelets but also on other time-frequency tools, namely the Hilbert-Huang Transform and the empirical mode decomposition.

with compression, but, instead, our main purpose is to analyze data to discover patterns or hidden information, then redundancy may be helpful, and the CWT can be used with advantage. With the advance of computer facilities and the availability of user-friendly toolboxes to compute the CWT, this transform will certainly become more and more popular among the economists.

Unfortunately, we do not find a good single reference for someone wanting to use (continuous) wavelet tools, such as: the continuous wavelet transform, the cross-wavelet transform, the wavelet coherency and the wavelet phase-difference. Not only the theoretic foundations are scattered among several papers and books, but also most codes freely available imply rigid assumptions, which do not give much freedom of choice to the researcher. For example, all the works cited in the previous paragraph make use of the Morlet wavelet and most of the codes do not provide many alternatives to the use of this particular wavelet.

All this helps to explain why continuous wavelets are not as popular in Economics as the discrete wavelet transform. Another possible explanation is that the techniques associated with DWT (and MODWT, for that matter) resemble more the traditional time series tools. E.g. Gallegati and Gallegati (2007) use MODWT to decompose the Industrial Production of the G-7 countries since 1961 in several scales (frequencies). Then, to analyze the evolution of the volatility of real economic activity, they estimate the wavelet variance at each scale, for each decade separately. This allows the researcher to use wavelets and, at the same time, have a foot on the traditional time series methods. As Yogo (2008) has shown, multiresolution wavelet analysis, which can very easily be performed with DWT, allows the decomposition of a time series into trend, cycle, and noise. The cyclical component of the wavelet-filtered series closely resembles the series filtered by the Baxter and King (1999) bandpass filter. More importantly, when a researcher uses the DWT he/she may have the feeling that multivariate tools are being used. For example, Ramsey and Lampart (1998a and 1998b) use DWT to individually analyze several time series (one each time), whose decompositions are then studied using traditional time-domain methods, such as Granger causality tests. However, in reality, wavelets were not used to do multivariate analysis.

Whitcher, Guttorp, Percival (2000), Gallegati (2008) — using the MODWT — and Crowley, Mayes and Maraun (2006), Crowley and Mayes (2008) and Aguiar-Conraria et al. (2008 and 2011a and b) — using CWT — have applied cross-wavelet analysis to uncover time–frequency interactions between two economic time series. Still, most surely, wavelets will not become very popular in economics until a concept analogous to the spectral partial-coherence is developed. This is one of the contributions of this paper: to develop the concepts of wavelet multiple coherency and wavelet partial coherency. We should note that Crowley, Mayes and Maraun (2006a and b) and Crowley (2007) have also some work on this direction. In their case they use multivariate spectral analysis using Hilbert wavelet pairs, originally proposed by Craigmile and Whitcher (2004).

This paper has five main purposes: (1) to give a self-contained summary on the most relevant theoretical results on continuous wavelet analysis, (2) to introduce the concepts of wavelet multiple coherency and wavelet partial coherency, (3) to introduce the economist to a new family of wavelets that have some desirable characteristics and that some authors believe to have the potential to become as popular as the Morlet wavelet, (4) to describe how the transforms can be implemented in practice, and (5) to provide a user-friendly Matlab toolbox implementing the referred wavelet tools, which the researcher can freely use and adapt to his/her own research.

This paper proceeds as follows. In Section 2, we describe the continuous wavelet transform. We also describe some of the characteristics of the Morlet wavelet, responsible for its popularity and a new class of analytic wavelets, the generalized Morse wavelets (GMWs), that can potentially become as useful as the Morlet wavelet.[3] In Section 3, we introduce the cross-wavelet tools, which include the cross-wavelet power, wavelet coherency and phase-difference. Section 4 extends the concepts of Section 3 to a multivariate setting. In each section, we construct an example to motivate and illustrate how to apply and interpret the described wavelet tools. Section 5 summarizes some of the results already obtained in terms of significance testing. In section 6, we provide three applications with real data. In one of them, we show evidence that corroborates the arguments of Blanchard and Simon (2001) about the Great Moderation, who have argued that the Great Moderation started well before 1983. In the second application, with the help of cross wavelet analysis, we study synchronism in international stock market returns. Finally, we will illustrate the usefulness of higher order wavelet tools, such as the partial wavelet coherencies and the partial phase-difference, to study the linkages between oil prices and the stock markets after controlling for some other macroeconomic variables. Section 7 concludes. In appendix A, we describe how to adapt the theory in order to implement it computationally and we include some proofs omitted in the main text. In appendix B, we give a brief description of some software packages and we describe our toolbox, the ASToolbox, as well as some of our computational choices.

2 Wavelet Analysis

2.1 Example 1: Motivation

We start with an example that applies the tools that will be explained later on. We realize that this is an unorthodox way to start a paper, but we hope that this will help to motivate the reader to invest some time in the rest of the paper.

The most common argument to justify the use of wavelet analysis over spectral analysis is the possibility of tracing transitional changes across time. It is as if we are estimating the power spectrum as a function of time. To illustrate this, consider the following highly stylized experiment with simulated data. We generate 50 years of monthly data according to the following data generating process:

$$y_t = \cos\left(\frac{2\pi}{p_1}t\right) + \cos\left(\frac{2\pi}{p_2}t\right), +\varepsilon_t \quad t = \frac{1}{12}, \frac{2}{12}, \ldots, 50, \tag{1}$$

where $p_1 = 10$ and $p_2 = 5$, if $20 \leq t \leq 30$, and $p_2 = 3$, otherwise.

Formula (1) tells us that the time series y_t is the sum of two periodic components.[4] The first periodic component represents a 10 year cycle, while the second periodic component shows some transient dynamics. In the beginning, it represents a 3 year cycle that, temporarily, changes to a 5 year cycle between the second and the third decades.

[3] We argue that while the Morlet wavelets represent the best compromise between frequency and time localization, GMWs allow for more flexibility, which can be useful if the researcher is specifically interested in having a better frequency or time localization.

[4] Although highly stylized, this formulation is not as restrictive as it may seem. An autoregressive process of order 2, or higher, with an oscillatory behavior, will have a solution that involves sines and cosines. We, therefore, could have generated similar time series using the more common auto regressive process. We chose to explicitly have a cosine

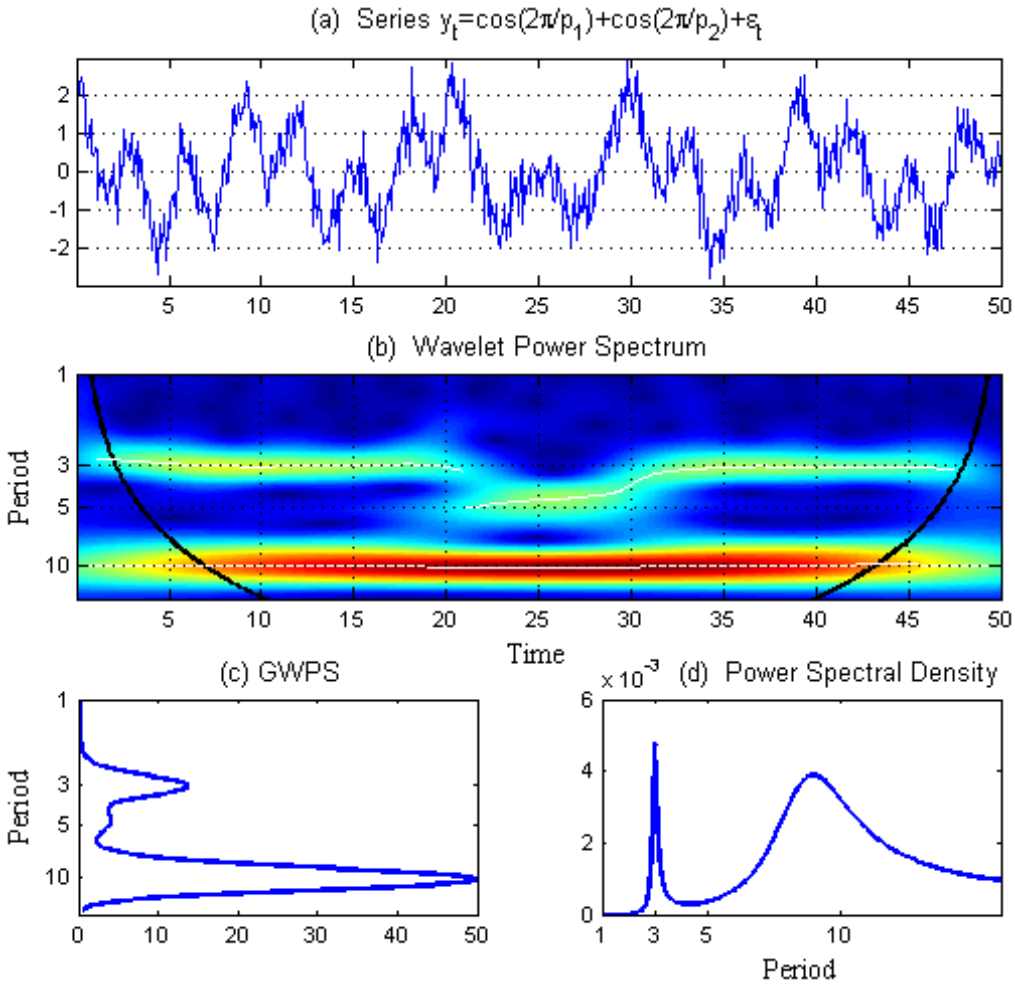

Figure 1: (a) $y_t = \cos(\frac{2\pi}{p_1}t) + \cos(\frac{2\pi}{p_2}t) + \varepsilon_t$. (b) Wavelet power spectrum of y_t.The cone of influence, which indicates the region affected by edge effects, is shown with a thick black line. The color code for power ranges from blue (low power) to red (high power). The white lines show the maxima of the undulations of the wavelet power spectrum. (c) Global wavelet power spectrum - average wavelet power for each frequency. (d) Fourier power spectral density.

This change in the dynamics is nearly impossible to spot in Figure 1 (a). Furthermore, if we use the traditional spectral analysis, the information on the transient dynamics is completely lost, as we can see in Figure 1 (d). The power spectral density estimate is able to capture both the 3-year and the 10-year cycles[5] but it completely fails to capture the 5-year cycle that occurred in the 20's. Comparing with Figure 1 (c), we observe that spectral analysis gives us essentially the same information as the global wavelet power spectrum, which is an average, across time, of the wavelet power spectrum.[6]

On the other hand, Figure 1 (b) shows the wavelet power spectrum itself. On the horizontal axis, we have the time dimension. The vertical axis gives us the periods. The power is given by the color. The color code for power ranges from blue (low power) to red (high power). Regions with

because the period of the oscillation is observed directly.

[5] We converted frequencies into period cycles.

[6] In all our examples and aplications, we use the Morlet Wavelet with $\omega_0 = 6$. See subsection 2.9.

warm colors represent areas of high power. The white lines show the maxima of the undulations of the wavelet power spectrum, therefore giving us an estimate of the cycle period. We observe a white line on period 10 across all times, meaning that there is a permanent cycle of this period. Both the red color and the black contour tell us that this cycle is strong and statistically significant. We are able to spot the three year period cycle that occurs between time zero and 20 and, again, between time 30 and 50. Finally, we are also able to spot a yellow region between time 20 and 30, with the white stripes identifying the cycle of period five. This means that a cycle of roughly 5-year periodicity, relatively important in explaining the total variance of the time series and taking place between year 20 and 30, was hidden by the Fourier power spectrum estimate.

Figure 1 (b) clearly illustrates the big advantage of wavelet analysis over spectral analysis. While the Fourier transform is silent about changes that happen across time, with wavelets we are able to estimate the power spectrum as a function of time and, therefore, we do not loose the time dimension. The wavelet power spectrum is able to capture not only the 3-year and 10-year cycles, but also to capture the change that occurred between years 20 and 30.

2.2 Notations and Conventions

In what follows, $L^2(\mathbb{R})$ denotes the set of square integrable functions, i.e. the set of functions defined on the real line and satisfying $\int_{-\infty}^{\infty} |x(t)|^2 dt < \infty$,with the usual inner product $\langle x, y\rangle := \int_{-\infty}^{\infty} x(t)y^*(t)dt$ and associated norm $\|x\| := \langle x, x\rangle^{\frac{1}{2}}$. Here, and in what follows, the asterisk superscript is used to denote complex conjugation and := means "by definition". Since the (squared) norm of $x(t)$, $\|x(t)\|^2 = \int_{-\infty}^{\infty} |x(t)|^2 dt$ is usually referred to as the *energy* of x, the space $L^2(\mathbb{R})$ is also known as the space of finite energy functions.

In these notes, we always use the convention $g(t) \leftrightarrow G(\omega)$ to denote a Fourier pair, i.e. we denote by the corresponding capital letter the Fourier transform of a given function. Hence, if $x(t) \in L^2(\mathbb{R})$, $X(\omega)$ will denote its Fourier transform, here defined as $X(\omega) := \int_{-\infty}^{\infty} x(t)e^{-i\omega t}dt$.[7]

The well-known Parseval relation, $\langle x(t), y(t)\rangle = \frac{1}{2\pi}\langle X(\omega), Y(\omega)\rangle$, valid for all $x(t), y(t) \in L^2(\mathbb{R})$, is an important result from which the Plancherel identity immediately follows $\|x(t)\|^2 = \frac{1}{2\pi}\|X(\omega)\|^2$.

2.3 Wavelets

The minimum requirement imposed on a function $\psi(t) \in L^2(\mathbb{R})$ to qualify for being a *mother* (*admissible* or *analyzing*) *wavelet* is that it satisfies a technical condition, usually referred to as the *admissibility condition*, which reads as follows:

$$0 < C_\psi := \int_{-\infty}^{\infty} \frac{|\Psi(\omega)|}{|\omega|} d\omega < \infty; \tag{2}$$

see Daubechies (1992). The constant C_ψ above is called the *admissibility constant*.

We should point out that the square integrability of $\psi(t)$ is a very mild decay condition and that, in practice, much more stringent conditions are imposed. In fact, for the purpose of providing a useful time-frequency localization, the wavelet must be a reasonable well localized function, both

[7] With the above convention of the Fourier Transform, ω is an *angular* (or radian) frequency. The relation to the more common Fourier frequency f is given by $f = \frac{\omega}{2\pi}$.

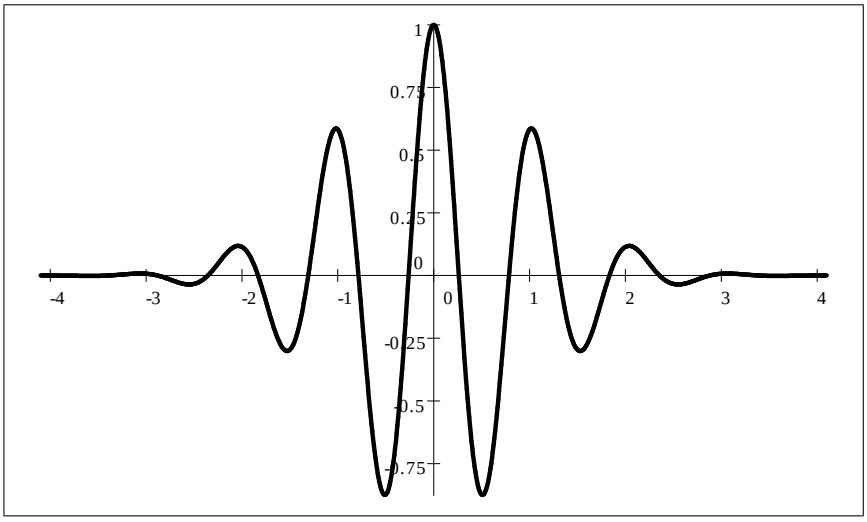

Figure 2: A typical wavelet

in the time domain as well as in the frequency domain. For functions with sufficient decay, it turns out that the admissibility condition (2) is equivalent to requiring that

$$\Psi(0) = \int_{-\infty}^{\infty} \psi(t)dt = 0; \tag{3}$$

again, see Daubechies (1992). This means that the function ψ has to wiggle up and down the $t-$axis, i.e. it must behave like a wave; this, together with the assumed decaying property, justifies the choice of the term wavelet (originally, in French, *ondelette*) to designate ψ — Figure 2. It is this property that allows for an effective localization in both time and frequency, contrary to the Fourier transform, which decomposes the signal in term of sines and cosines, i.e. infinite duration waves.

2.4 The Continuous Wavelet Transform

Starting with a mother wavelet ψ, a family $\psi_{\tau,s}$ of "wavelet daughters" can be obtained by simply scaling and translating ψ:

$$\psi_{\tau,s}(t) := \frac{1}{\sqrt{|s|}}\psi\left(\frac{t-\tau}{s}\right), \quad s, \tau \in \mathbb{R},\, s \neq 0, \tag{4}$$

where s is a scaling or dilation factor that controls the width of the wavelet and τ is a translation parameter controlling the location of the wavelet. Scaling a wavelet simply means stretching it (if $|s| > 1$) or compressing it (if $|s| < 1$), while translating it simply means shifting its position in time.

Given a time series $x(t) \in L^2(\mathbb{R})$, its *continuous wavelet transform* (CWT) with respect to the wavelet ψ is a function of two variables, $W_{x;\psi}(\tau, s)$:

$$W_{x;\psi}(\tau, s) = \int_{-\infty}^{\infty} x(t)\frac{1}{\sqrt{|s|}}\psi^*\left(\frac{t-\tau}{s}\right)dt. \tag{5}$$

The position of the wavelet in the time domain is given by τ, while its position in the frequency

domain is given by s (the relation between s and the frequency will soon be clearer). Therefore the wavelet transform, by mapping the original series into a function of τ and s, gives us information simultaneously on time and frequency. The formulas of the wavelet and the Fourier transforms are very similar. The main differences are that in the Fourier transform we have no time localization parameter and that we have cosine and sine functions instead of a wavelet function.

By the well-known properties of the Fourier transform, one immediately sees that the CWT (5) may also be represented in the frequency, as

$$W_x(\tau, s) = \frac{\sqrt{|s|}}{2\pi} \int_{-\infty}^{\infty} \Psi^*(s\omega) X(\omega) e^{i\omega\tau} d\omega. \text{ [8]} \tag{6}$$

2.5 Inversion of CWT

The importance of the admissibility condition (2) comes from the fact that its fulfilment guarantees that the energy of the original function $x(t)$ is preserved by the wavelet transform, i.e., the following Parseval-type relation holds: $\int_{-\infty}^{\infty} |x(t)|^2 dt = \frac{1}{C_\psi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |W_x(\tau, s)|^2 \frac{d\tau ds}{s^2}$,which, in turn, ensures the possibility of recovering $x(t)$ from its wavelet transform. In fact, due to the high redundancy of this transform (note that a function of one variable is mapped into a bivariate function), many reconstruction formulas are available. For example, when the wavelet ψ is real-valued, it is possible to reconstruct $x(t)$ by using the formula

$$x(t) = \frac{2}{C_\psi} \int_0^{\infty} \left[\int_{-\infty}^{\infty} W_x(\tau, s) \psi_{\tau,s}(t) d\tau \right] \frac{ds}{s^2}, \tag{7}$$

showing that no information is lost if we restrict the computation of the transform only to positive values of the scaling parameter s, which is a usual requirement, in practice.

One can also limit the integration over a selected range of scales, performing a band-pass filtering of the original series. To our knowledge, except in Aguiar-Conraria et al. (2008) and Aguiar-Conraria and Soares (2011b) nobody has used the inversion formula as a band-pass filter. Our experience tell us that not much is gained when compared to the more common band-pass filters, such as the Baxter and King (1999) and Christiano and Fitzgerald (2003) filters.

2.6 Wavelet Power Spectrum and Wavelet Phase

In analogy with the terminology used in the Fourier case, the (local) *wavelet power spectrum* (sometimes called *scalogram* or *wavelet periodogram*) is defined as

$$(WPS)_x(\tau, s) = |W_x(\tau, s)|^2 . \tag{8}$$

The wavelet power spectrum may be averaged over time for comparison with classical spectral methods. When the average is taken over all times, we obtain the *global wavelet power spectrum*:

$$(GWPS)_x(s) = \int_{-\infty}^{\infty} |W_x(\tau, s)|^2 d\tau. \tag{9}$$

[8] When the wavelet ψ is implicit from the context, we abbreviate the notation and simply write W_x for $W_{x;\psi}$.

When the wavelet ψ is complex-valued, the corresponding wavelet transform $W_x(\tau, s)$ is also complex-valued. In this case, the transform can be separated into its real part, $\Re\{W_x(\tau, s)\}$, and imaginary part, $\Im\{W_x(\tau, s)\}$, or in its *amplitude*, $|W_x(\tau, s)|$, and *phase* (or phase-angle), $\phi_x(\tau, s) : W_x(\tau, s) = |W_x(\tau, s)|\, e^{i\phi_x(\tau,s)}$. The phase-angle $\phi_x(\tau, s)$ of the complex number $W_x(\tau, s)$ can be obtained from the formula:

$$\phi_x(\tau, s) = \text{Arctan } \left(\frac{\Im\{W_x(\tau, s)\}}{\Re\{W_x(\tau, s)\}} \right). \text{ [9]}$$

For real-valued wavelet functions, the imaginary part is constantly zero and the phase is, therefore, undefined. Hence, in order to separate the phase and amplitude information of a time series, it is important to make use of complex wavelets. In this case, it is convenient to choose a wavelet $\psi(t)$ whose Fourier transform is supported on the positive real-axis only, i.e. is such that $\Psi(\omega) = 0$ for $\omega < 0$. A wavelet satisfying this property is called *analytic* or *progressive*. When ψ is analytic and $x(t)$ is real, reconstruction formulas involving only positive values of the scale parameter s are still available; in particular, if the wavelet satisfies $0 < |K_\psi| < \infty$, where $K_\psi := \int_0^\infty \frac{\Psi^*(\omega)}{\omega} d\omega$, then one can use the following reconstruction formula, known as the *Morlet formula*, which is particularly useful for numerical applications:

$$x(t) = 2\Re \left[\frac{1}{K_\psi} \int_0^\infty W_x(t, s) \frac{ds}{s^{3/2}} \right]; \tag{10}$$

see, e.g. Farge (1992) or Holschneider (1995).

When the wavelet ψ is analytic, the corresponding wavelet transform is called an *analytic wavelet transform* (AWT).

Remark 1 *Throughout the rest of this paper, we assume that all the wavelets considered are analytic and hence, that the wavelet transform is computed only for positive values of the scaling parameter s. For this reason, in all the formulas that would normally involve the quantity $|s|$, this will be replaced by s.*

2.7 Localization Properties

The Heisenberg uncertainty principle was derived in quantum mechanics. It basically stated that certain pairs of physical properties, such as position and momentum, cannot be simultaneously known to arbitrarily high precision: The more precisely one property is measured, the less precisely the other can be measured. Applied in our context, if we want precision in the frequency we have give up some precision in time. The Fourier transform does that: it has an excellent frequency localization but the time information is lost. With wavelets, information on both is kept. In this subsection, we will exactly describe this trade off.

[9] We use Arctan to denote the following extension of the usual principal component of the arctan function (whose range is $(-\pi/2, \pi/2)$): $\text{Arctan } \left(\frac{b}{a}\right) = \begin{cases} \text{Arctan } (\frac{b}{a}) & a > 0, \\ \text{Arctan } (\frac{b}{a}) + \pi & a < 0, \quad b \geq 0, \\ \text{Arctan } (\frac{b}{a}) - \pi & a < 0, \quad b < 0, \\ \pi/2 & a = 0, \quad b \geq 0, \\ -\pi/2 & a = 0, \quad b < 0. \end{cases}$

In order to describe the time-frequency localization properties of the CWT, we have to assume that both the wavelet $\psi(t)$ and its Fourier transform $\Psi(\omega)$ are well localized functions. More precisely, these functions must have sufficient decay to guarantee that the quantities defined below are all finite.[10] We define the *center in time* of the wavelet ψ , $\mu_{t;\psi}$, by

$$\mu_{t;\psi} = \frac{1}{\|\psi\|^2} \int_{-\infty}^{\infty} t \, |\psi(t)|^2 \, dt,$$

and, as a measure of concentration of ψ around its center, we take the *standard deviation in time* (also known as the *radius in time*), $\sigma_{t;\psi} = \frac{1}{\|\psi\|} \left\{ \int_{-\infty}^{\infty} \left(t - \mu_{t;\psi}\right)^2 |\psi(t)|^2 \, dt \right\}^{\frac{1}{2}}$. The *center in frequency*, $\mu_{\omega;\psi}$, and the *standard deviation (or radius) in frequency*, $\sigma_{\omega;\psi}$, are defined in a similar way:

$$\mu_{\omega;\psi} = \frac{1}{\|\Psi\|^2} \int_{-\infty}^{\infty} \omega \, |\Psi(\omega)|^2 \, d\omega \tag{11}$$

and $\sigma_{\omega;\psi} = \frac{1}{\|\Psi\|} \left\{ \int_{-\infty}^{\infty} \left(\omega - \mu_{\omega;\psi}\right)^2 |\Psi(\omega)|^2 \, d\omega \right\}^{\frac{1}{2}}$.[11]

The quantities μ_t and σ_t are the mean and standard deviation of the probability density function (p.d.f.) defined by $|\psi(t)|^2/\|\psi\|^2$. The same is true for μ_ω, σ_ω and $|\Psi(\omega)|^2/\|\Psi\|^2$. Therefore, it should not come as a surprise that the interval $[\mu_t - \sigma_t, \mu_t + \sigma_t]$ is the set where $\psi(t)$ attains its "most significant" values whilst the interval $[\mu_\omega - \sigma_\omega, \mu_\omega + \sigma_\omega]$ plays the same role for $\Psi(\omega)$. The rectangle

$$H_\psi := [\mu_t - \sigma_t, \mu_t + \sigma_t] \times [\mu_\omega - \sigma_\omega, \mu_\omega + \sigma_\omega] \tag{12}$$

in the (t, ω) −plane is called the *Heisenberg box* or *window* for the function ψ. We then say that ψ is localized around the point (μ_t, μ_ω) of the time-frequency plane, with *uncertainty* given by $\sigma_t \sigma_\omega$. In our context, the Heisenberg uncertainty principle establishes the following lower bound:

$$\sigma_t \sigma_\omega \geq \frac{1}{2}. \tag{13}$$

Recalling that the wavelet daughter $\psi_{\tau,s}$ is obtained from its mother ψ by a simple translation by τ and a scaling by s, it is easy to show that the center and radius in time of $\psi_{\tau,s}$ are given by $\mu_{t;\psi_{\tau,s}} = \tau + s\mu_t$ and $\sigma_{t;\psi_{\tau,s}} = s\sigma_t$ and that the center and radius in frequency of $\psi_{\tau,s}$ are given by $\mu_{\omega;\psi_{\tau,s}} = \frac{\mu_\omega}{s}$ and $\sigma_{\omega;\psi_{\tau,s}} = \frac{\sigma_\omega}{s}$. In particular, if the mother wavelet ψ is centered at $t = 0$, i.e. if $\mu_t = 0$,[12] then the window associated with $\psi_{\tau,s}$ becomes

$$H_{\psi_{\tau,s}} = [\tau - s\sigma_t, \tau + s\sigma_t] \times \left[\frac{\mu_\omega}{s} - \frac{\sigma_\omega}{s}, \frac{\mu_\omega}{s} + \frac{\sigma_\omega}{s} \right]. \tag{14}$$

In this case, one has that $W_x(\tau, s) \approx \int_{\tau - s\sigma_t}^{\tau + s\sigma_t} x(t) \psi_{\tau,s}^*(t) dt$ and, by the Parseval relation, $W_x(\tau, s) \approx 2\pi \int_{\frac{\mu_\omega}{s} - \frac{\sigma_\omega}{s}}^{\frac{\mu_\omega}{s} + \frac{\sigma_\omega}{s}} X(\omega) \Psi_{s,\tau}(\omega) d\omega$.

We thus conclude that the continuous wavelet transform $W_x(\tau, s)$ gives us temporal information on $x(t)$ around the instant $t(\tau) = \tau$, with precision $s\sigma_t$, and frequency information about $X(\omega)$

[10] The precise requirements are that $|\psi(t)| < C(1 + |t|)^{-(1+\epsilon)}$ and $|\Psi(\omega)| < C(1 + |\omega|)^{-(1+\epsilon)}$, for $C < \infty$, $\epsilon > 0$.

[11] If the wavelet ψ is known from the context, we will suppress the index ψ in the notation of the above quantities, e.g. we will simply use μ_t for $\mu_{t;\psi}$, etc.

[12] Note that this can easily be achieved by an appropriate translation.

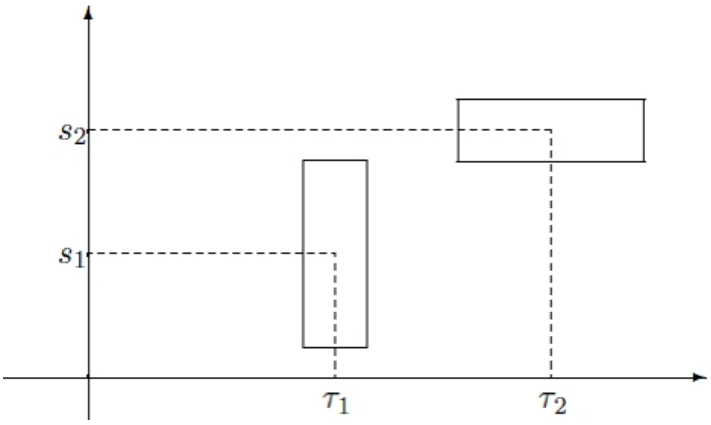

Figure 3: Windows associated with a Continuous Wavelet Transform

around the frequency

$$\omega(s) = \frac{\mu_\omega}{s}, \tag{15}$$

with precision $\frac{\sigma_\omega}{s}$.

Although the area of the windows are constant (given by $4\sigma_t\sigma_\omega$), their dimensions change according to the scale; the windows stretch for large values of s (broad scales s – low frequencies $\omega_s = \frac{\sigma_\omega}{s}$) and compress for small values of s (fine scale – high frequencies $\frac{\sigma_\omega}{s}$). This is what some authors mean when they say that one major advantage afforded by the wavelet transform — when compared with the most famous alternative, the *short time or windowed Fourier transform* — is its ability to adjust to the local analysis of a time series in the sense that the length of the wavelets varies endogenously: it stretches into a long wavelet function to measure the low frequency movements; and it compresses into a short wavelet function to measure the high frequency movements.

2.8 Scale/Frequency Relation and Fourier Factor

Strictly speaking, the wavelet transform provides us a time-scale representation of the function being analyzed and not a time-frequency representation. Formula (15) is commonly used to convert scales into frequencies. However, we should have in mind that this inverse relation between scale and frequency corresponds to a particular interpretation and that there are other meaningful ways of assigning frequencies to scales. As Meyers et al. (1993) say, "for a general wavelet, the relation between scale and the more common Fourier wavelength is not necessarily straightforward; for example, some wavelets are highly irregular without any dominant periodic components. In those cases it is probably a meaningless exercise to find a relation between the two disparate measures of distance." This aspect, as we explain later, makes the choice of the particular wavelet an important option.

There are at least three different meaningful ways to convert scales into frequencies. In Lilly and Olhede (2009), the authors consider, apart from the value of $\mu_{\omega;\psi}$ given by (11), which they call the *energy-frequency*, and which, for convenience, we will now denote by ω_ψ^E, two other specific frequencies associated with the wavelet: the *peak frequency*, ω_ψ^P, defined as the frequency at which the magnitude of the Fourier transform of ψ, $|\Psi(\omega)|$, is maximized, i.e. $|\Psi(\omega_\psi^P)| = \sup_{\omega\in\mathbb{R}} |\Psi(\omega)|$;

and the *central instantaneous frequency*, ω_ψ^I, defined as the value that the *time-varying instantaneous frequency* of the wavelet takes at its center (here assumed to be 0), i.e. $\omega_\psi^I = \breve{\omega}_\psi(0)$,where $\breve{\omega}(t)$, is the time-varying instantaneous frequency of the wavelet, defined by $\breve{\omega}_\psi(t) = \frac{d}{dt}\Im\{\ln\psi(t)\}$. To each of the three specific frequencies, $\omega_\psi^E, \omega_\psi^P$ and ω_ψ^I they associate an interpretation of scale as frequency. More precisely, they define

$$\omega(s) = \frac{\omega_\psi}{s}, \tag{16}$$

with ω_ψ denoting any of the three specific frequencies. Note that $\omega(s)$, as well as $\omega_\psi^E, \omega_\psi^P$ and ω_ψ^I, are all angular frequencies. If we prefer a relation between the scale and the usual "Fourier" frequency f (expressed in cycles per unit time), we have

$$f(s) = \frac{\omega_\psi}{2\pi s}. \tag{17}$$

$\mathrm{Ff} = \frac{2\pi}{\omega_\psi}$ is called *Fourier factor* of the wavelet and is used, in the programs, to convert scales to periods.

To see that the three correspondences $\omega(s) = \frac{\omega_\psi}{s}$, $\omega_\psi = \omega_\psi^P, \omega_\psi^E, \omega_\psi^I$ are all meaningful (although in different senses) we refer the reader to the mentioned paper by Lilly and Olhede (2009).

Naturally, it will be convenient to choose a wavelet whose associated frequencies $\omega_\psi^E, \omega_\psi^P$ and ω_ψ^I have all the same (or, at least, very similar) value, since this will give us a unified view of the relation between frequency and scale.

2.9 Analytic Wavelets

The admissibility condition (2) is a very weak condition and, in theory, there are infinitely many wavelets. In practice, which wavelet to use is an important aspect to be taken into account, and will be dictated by the kind of application one has in mind. To study the synchronism between different time series, it is important to select a wavelet whose corresponding transform contains information on both amplitude and phase, and hence, a complex-valued analytic wavelet is a natural choice.

As stated in Lilly and Olhede (2009), the analytic wavelet transform (AWT) is the basis for the *wavelet ridge* method, which recovers time-varying estimates of instantaneous amplitude, phase, and frequency of a modulated oscillatory signal from the time/scale plane (Delprat et al. 1992 and Mallat 1998). On the other hand, the analytic wavelet transform can also be useful for application to very time-localized structures Tu, Hwang and Ho (2005). The many useful features of analytic wavelets are covered in more depth by Selesnick, Baraniuk and Kinsbury (2005); see also Olhede and Walden (2002) and Lilly and Olhede (2009).

In this section, we summarize some properties that explain why the Morlet wavelet is the most used wavelet in practice (to our knowledge, every application of the continuous wavelet transform in Economics has used this choice). We also present a particularly important family of analytic wavelets, the generalized Morse wavelets (GMWs), which are more flexible and can be used as an alternative to the Morlet wavelet when one prefers better time localized or frequency localized wavelets. The results of this section are derived from the papers by Olhede and Walden (2002) and Lilly and Olhede (2009 and 2010).

2.9.1 Morlet Wavelets

The *Morlet wavelets* are a one-parameter family of functions, first introduced in Goupillaud, Grossman and Morlet (1984), and given by

$$\psi_{\omega_0}(t) = Ke^{i\omega_0 t}e^{-\frac{t^2}{2}}. \tag{18}$$

Strictly speaking, the above functions are not true wavelets, since they fail to satisfy the admissibility condition.[13] For $\psi_{\omega_0}(t)$ to have unit energy, the normalizing constant K must be chosen as

$$K = \pi^{-1/4}, \tag{19}$$

which, from now on, we will always assume to be true. The Fourier transform of the normalized wavelet is given by

$$\Psi_{\omega_0}(\omega) = \sqrt{2}\pi^{1/4}\, e^{-\frac{1}{2}(\omega-\omega_0)^2} \tag{20}$$

and, hence, $\Psi_{\omega_0}(0) = \sqrt{2}\pi^{1/4}\, e^{-{\omega_0}^2/2} \neq 0$. However, for sufficiently large ω_0, e.g. $\omega_0 > 5$, the values of $\Psi\omega_0(\omega)$ for $\omega \leq 0$ are so small that, for numerical purposes, Ψ_{ω_0} can be considered as an analytic wavelet; see Foufoula-Georgiou and Kumar (1994).

The Morlet wavelet became the most popular of the complex valued wavelets mainly because of four interesting properties. First, for numerical purposes, as we have just seen, it can be treated as an analytic wavelet. Second, the peak frequency, the energy frequency and the central instantaneous frequency of the Morlet wavelet are all equal and given by

$$\omega^P_{\psi_{\omega_0}} = \omega^E_{\psi_{\omega_0}} = \omega^I_{\psi_{\omega_0}} = \omega_0, \tag{21}$$

facilitating the conversion from scales to frequencies. Using formula (17), and for the most common choice choice of $\omega_0 = 6$, we have that $f = \frac{6}{2\pi s} \approx \frac{1}{s}$. Third, the Heisenberg box area reaches its lower bound with this wavelet, i.e. the uncertainty attains the minimum possible value: $\sigma_{t;\psi_{\omega_0}}\sigma_{\omega;\psi_{\omega_0}} = \frac{1}{2}$. In this sense, the Morlet wavelet has optimal joint time-frequency concentration. Finally, the time radius and the frequency radius are equal, $\sigma_{t;\psi_{\omega_0}} = \sigma_{\omega;\psi_{\omega_0}} = \frac{1}{\sqrt{2}}$,and, therefore, this wavelet represents the best compromise between time and frequency concentration. To our knowledge, at least in economics, every paper uses some value of $\omega_0 \in [5, 6]$.

2.9.2 Generalized Morse Wavelets

In spite of its usefulness, the Morlet wavelet is not very versatile, because it depends on just one parameter. If one is interested in having a better localization in frequency (or in time) one cannot adjust the Morlet wavelet. On the other hand, it is quite common to have referee reports asking for the robustness of the results to the choice of the Morlet wavelet (even if you are very careful at explaining the optimal characteristics of the Morlet wavelet).

Finally, although it is true that the Morlet wavelet has optimal joint time-frequency concentration in the Heisenberg sense, it is also true that other criteria are available. "The whole set of generalized

[13] In order to fulfill the admissibility condition, a correction term has to be added, as: $\psi_{\omega_0}(t) = K\left(e^{i\omega_0 t} - e^{-\omega_0^2/2}\right)e^{-t^2/2}$.

Morse wavelets are optimally localized in that they maximize the eigenvalues of a joint time-frequency localization operator (...) and indeed this is the way the generalized Morse wavelets were initially constructed." — in Lilly and Olhede (2009).

The generalized Morse wavelets (GMWs) are a two-parameter family of wavelets, defined, in the frequency domain, by

$$\Psi_{\beta,\gamma}(\omega) = K_{\beta,\gamma}\, H(\omega)\, \omega^{\beta} e^{-\omega^{\gamma}} \tag{22}$$

where $K_{\beta,\gamma}$ is a normalizing constant and $H(\omega)$ is the Heaviside unit step function.

To be a valid wavelet, one must have $\beta > 0$ and $\gamma > 0$. By varying these two parameters, the generalized Morse wavelets can be given a broad range of characteristics while remaining exactly analytic. In fact, these wavelets form a very wide family that subsumes many other types of wavelets. Lilly and Olhede showed that the generalized Morse wavelets encompass two other popular families of analytic wavelets: the Cauchy or Klauder wavelet family (for $\gamma = 1$), the Paul wavelets (which correspond to the case $\gamma = 1$ and $\beta \in \mathbb{N}$) and the analytic "Derivative of Gaussian" (DOG) wavelets (for $\gamma = 2$).

Unfortunately, contrary to the Morlet case, for the generalized Morse wavelets, we do not have a single way to convert scales into frequencies. This is so because the peak frequency, $\omega^{P}_{\beta,\gamma} = \left(\frac{\beta}{\gamma}\right)^{1/\gamma}$, is different from the energy frequency, $\omega^{E}_{\beta,\gamma} = \frac{1}{2^{1/\gamma}} \frac{\Gamma(\frac{2\beta+2}{\gamma})}{\Gamma(\frac{2\beta+1}{\gamma})}$, which is different from the central instantaneous frequency, $\omega^{I}_{\beta,\gamma} = \frac{\Gamma(\frac{\beta+2}{\gamma})}{\Gamma(\frac{\beta+1}{\gamma})} = 2^{1/\gamma}\omega^{E}_{\beta/2,\gamma}$. For the economist, used to think about frequencies, this is an obvious disadvantage of this family of wavelets.

2.9.3 Measures for the Morlet and the Generalized Morse Wavelets

A table with the localization measures (time-radius, frequency-radius and Heisenberg area) for GMW $\psi_{\beta;\gamma}$, for some values of β and γ is given below. In the last row, we also indicate the measures for the Morlet wavelet.[14]

As we have said before, in economics every single application of the CWT that we know of uses the Morlet with $\omega_0 \in [5, 6]$. And there are good reasons for this choice, as we can in Table 1. However, if for a particular application the researcher needs to have a better frequency localization, then the GMW provides a good alternative, because it is flexible while remaining exactly analytic. For example, for $\beta = 10$ and $\gamma = 10$ the Heisenberg uncertainty (0.529) is close to its lower bound and the frequency accuracy is very good: $\sigma_f = 0.075$ (with $\sigma_t = 7.1$). If instead the researcher wants a very well time localized wavelet, then, with $\beta = 10$ and $\gamma = 1/2$, one would have good localization in time, $\sigma_t = 0.004$, bad localization in frequency, $\sigma_f = 140$, and the Heisenberg uncertainty, 0.54, is not far from its lower bound.

The best compromise is achieved by the Morlet wavelet, however if one wants to explore the trade off between time and frequency precision then, probably, the generalized Morse wavelets are the best wavelet function available.

[14] In our toolbox we provide a code that allows the researcher to compute these measures for any combination of β and γ. The formulas for the radius in time and in frequency, needed to compute the Heisenberg uncertainty, are given in Lilly and Olhede (2009). The same code may be used to calculate $\omega^{P}_{\beta,\gamma}$, $\omega^{E}_{\beta,\gamma}$, $\omega^{I}_{\beta,\gamma}$.

β	γ	σ_t	σ_f	Heisenberg Uncertainty
1	1/2	0.224	8.874	1.984
	1	1.000	0.866	0.866
	2	1.732	0.337	0.583
	3	2.062	0.258	0.531
	5	2.470	0.214	0.528
	7	2.779	0.200	0.556
	10	3.170	0.193	0.611
2	1/2	0.058	17.78	1.023
	1	0.577	1.118	0.646
	2	1.528	0.344	0.525
	3	2.097	0.241	0.504
	5	2.804	0.183	0.513
	7	3.296	0.164	0.540
	10	3.876	0.152	0.588
3	1/2	0.029	28.53	0.815
	1	0.447	1.323	0.592
	2	1.483	0.347	0.514
	3	2.194	0.229	0.501
	5	3.109	0.164	0.511
	7	3.742	0.142	0.533
	10	4.472	0.128	0.572
5	1/2	0.012	54.52	0.675
	1	0.333	1.658	0.553
	2	1.453	0.349	0.508
	3	2.354	0.213	0.500
	5	3.582	0.142	0.508
	7	4.439	0.118	0.523
	10	5.418	0.102	0.552
7	1/2	0.007	85.58	0.621
	1	0.277	1.937	0.537
	2	1.441	0.351	0.505
	3	2.475	0.202	0.500
	5	3.945	0.128	0.506
	7	4.986	0.104	0.518
	10	6.172	0.087	0.539
10	1/2	0.004	140.1	0.583
	1	0.229	2.291	0.526
	2	1.433	0.351	0.503
	3	2.616	0.191	0.500
	5	4.378	0.115	0.504
	7	5.648	0.091	0.513
	10	7.100	0.075	0.529

Morlet	σ_t	σ_f	Heisenberg Uncertainty
	0.707	0.707	0.500

Table 1: Measures for some members of the familly of generalized Morse wavelets

3 Cross-Wavelet Analysis

In many applications, one is interested in detecting and quantifying relationships between two non-stationary time series. The concepts of cross-wavelet power, wavelet coherency and wavelet phase-difference are natural generalizations of the basic wavelet analysis tools that enable us to appropriately deal with the time-frequency dependencies between two time series.

Remark 2 *From now on, all the quantities we are going to introduce (e.g. cross-wavelet transform, wavelet coherency, etc.) are functions of time and scale (or frequency). In order to simplify the notation, we will describe these quantities for a specific value of the argument* (τ, s) *and this value of the argument will, unless strictly necessary, be omitted in the formulas.*

3.1 Cross-Wavelet Transform and Cross-Wavelet Power

The *cross-wavelet transform* (XWT) of two time series, $x(t)$ and $y(t)$, first introduced by Hudgins, Friehe and Mayer (1993), is simply defined as

$$W_{xy} = W_x W_y^*, \tag{23}$$

where W_x and W_y are the wavelet transforms of x and y, respectively.[15] The *cross-wavelet power* is

$$(XWP)_{xy} = |W_{xy}| . \tag{24}$$

While we can interpret the wavelet power spectrum as depicting the local variance of a time series, the cross-wavelet power of two time series depicts the local covariance between these time series at each time and frequency. Therefore, the cross-wavelet power gives us a quantified indication of the similarity of power between two time series.

3.2 Complex Wavelet Coherency

In analogy with the concept of coherency used in Fourier analysis, given two time series $x(t)$ and $y(t)$ one can define their *complex wavelet coherency* ϱ_{xy} by:

$$\varrho_{xy} = \frac{S\left(W_{xy}\right)}{\left[S\left(|W_x|^2\right) S\left(|W_y|^2\right)\right]^{1/2}}, \tag{25}$$

where S denotes a smoothing operator in both time and scale; smoothing is necessary, because, otherwise, coherency would be identically one at all scales and times.[16] Time and scale smoothing can be achieved by convolution with appropriate windows; see Cazelles, Chavez, de Magny, Guégan and Hales (2007) or Grinsted, Moore and Jevrejeva (2004), for details.

3.3 Wavelet Coherency and Phase-Difference

The complex wavelet coherency can be written in polar form, as $\varrho_{xy} = \left|\varrho_{xy}\right| e^{i\phi_{xy}}$. The absolute value of the complex wavelet coherency is called the *wavelet coherency* and is denoted by R_{xy}, i.e.

$$R_{xy} = \frac{\left|S\left(W_{xy}\right)\right|}{\left[S\left(|W_x|^2\right) S\left(|W_y|^2\right)\right]^{1/2}}, \tag{26}$$

with $0 \leq R_{xy}(\tau, s) \leq 1$.[17] The angle ϕ_{xy} of the complex coherency is called the *phase-difference* (phase lead of x over y), i.e.

$$\phi_{xy} = \text{Arctan}\left(\frac{\Im\left(S\left(W_{xy}\right)\right)}{\Re\left(S\left(W_{xy}\right)\right)}\right) \tag{27}$$

Some authors prefer a slightly different definition, $\text{Arctan}\left(\frac{\Im(W_{xy})}{\Re(W_{xy})}\right)$. In this case, one has $\phi_{xy} = \phi_x - \phi_y$,[18] hence the name phase-difference. The advantage of this definition is that because the phase-difference is not affected by the smoothing choice, it is fully consistent with the individual phases.

A phase-difference of zero indicates that the time series move together at the specified time-frequency; if $\phi_{xy} \in (0, \frac{\pi}{2})$, then the series move in phase, but the time series x leads y; if $\phi_{xy} \in (-\frac{\pi}{2}, 0)$, then it is y that is leading; a phase-difference of π (or $-\pi$) indicates an anti-phase relation; if

[15] When $y = x$, we obtain the Wavelet Power Spectrum $W_{xx} = |W_x|^2 = (WPS)_x$.
[16] The same happens with the Fourier coherency.
[17] At points (τ, s) for which $S\left(|W_x(\tau, s)|^2\right) S\left(|W_y(\tau, s)|^2\right) = 0$ we define $R_{xy}(\tau, s) = 0$.
[18] To be more precise, the above relation holds after we convert $\phi_x - \phi_y$ into an angle in the interval $[-\pi, \pi]$.

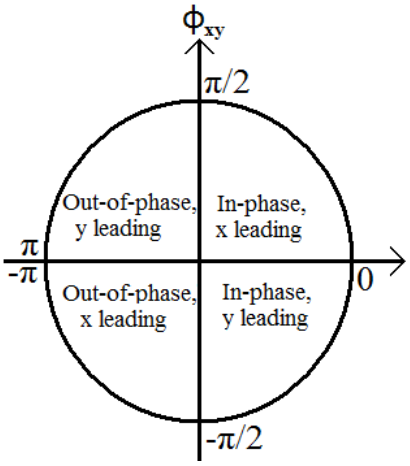

Figure 4: Phase-difference circle

$\phi_{xy} \in (\frac{\pi}{2}, \pi)$, then y is leading; time series x is leading if $\phi_{xy} \in (-\pi, -\frac{\pi}{2})$. This is illustrated in Figure 4.

One can easily convert the phase-difference into the *instantaneous time-lag* between the two time series x and y:

$$(\Delta T)_{xy}(\tau, s) = \frac{\phi_{xy}(\tau, s)}{\omega(s)}, \tag{28}$$

where $\omega(s)$ is the angular frequency that corresponds to the scale s.

3.4 Example 2: The Cross-Wavelet and the Phase-Difference

Consider now two time series that share two common cycles, with some delays:

$$x_t = \sin\left(\frac{2\pi}{3}t\right) + 3\sin\left(\frac{2\pi}{6}t\right) + \varepsilon_{x,t}, \qquad t = 0, \frac{1}{12}, \frac{2}{12}, \ldots, 50, \tag{29}$$

$$y_t = \begin{cases} 4\sin\left(\frac{2\pi}{3}(t+\frac{5}{12})\right) - 3\sin\left(\frac{2\pi}{6}(t-\frac{10}{12})\right) + \varepsilon_{y,t}, & t = 0, \frac{1}{12}, \frac{2}{12}, \ldots, 25, \\ 4\sin\left(\frac{2\pi}{3}(t-\frac{5}{12})\right) - 3\sin\left(\frac{2\pi}{6}(t+\frac{10}{12})\right) + \varepsilon_{y,t}, & t = 25 + \frac{1}{12}, 25 + \frac{2}{12}, \ldots, 50; \end{cases} \tag{30}$$

see Figure 5 (a). Looking at the formulas, it is clear that x_t and y_t share 3-year and 6-year cycles. However, how their cycles relate to each other evolves with time and is different across frequencies. Consider the shorter period cycle, the 3-year cycle. The cycles are positively correlated. For the first half of the sample, the y_t cycle precedes the x_t cycle by 5 months; in the second half of the sample the y_t cycle lags the x_t cycle.

These features are captured in Figure 5 (b)-(d). On the left, we have the wavelet coherency. On the right we have the phases and phase-difference computed for two different frequency bands. On the top, we compute the phases for the $2.5 \sim 3.5$ year frequency band. In the bottom, we consider the $5 \sim 7$ year frequency band. The green line represents the y_t phase and the blue represents the x_t phase. The red line represents the phase-difference between y_t and x_t.

That both series have common and highly correlated 3-year and 6-year cycles is revealed by

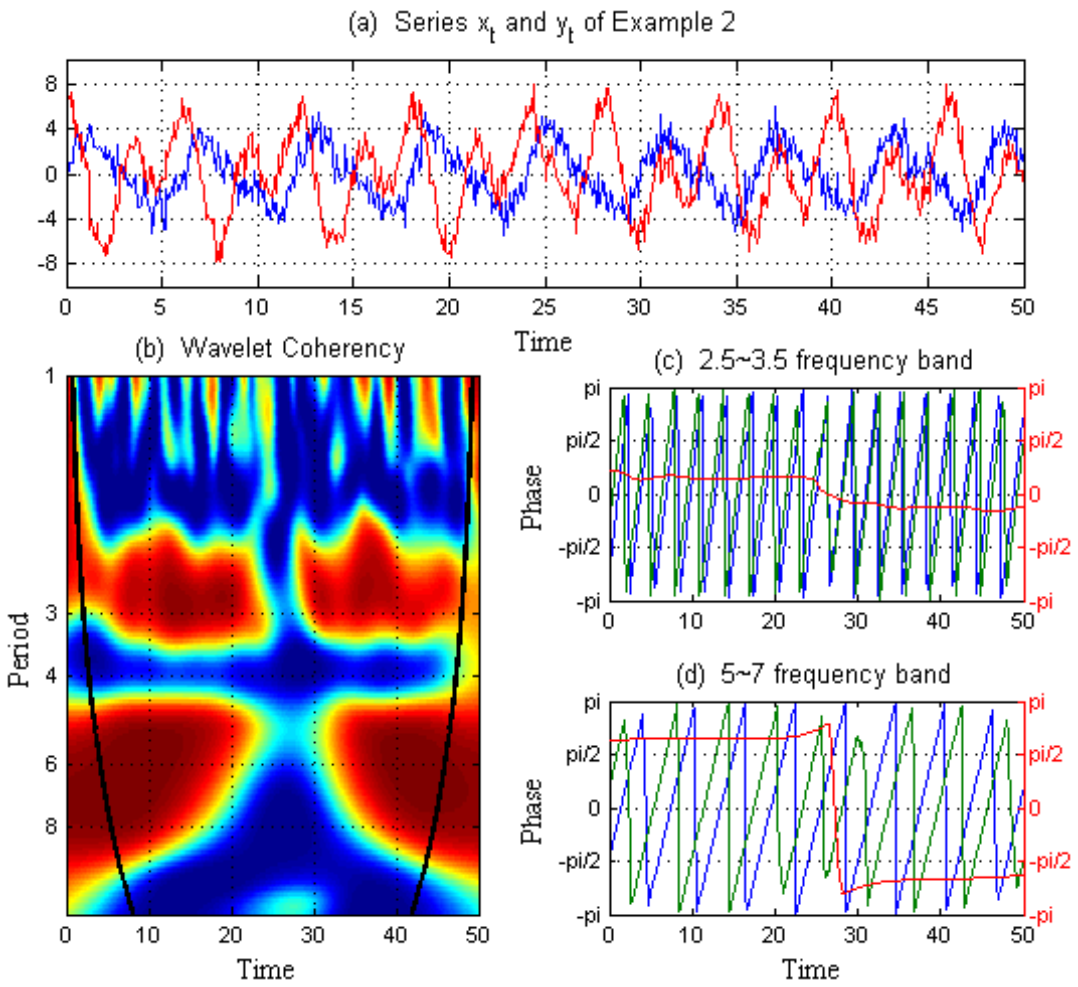

Figure 5: (a) x_t (blue line) given by Eq.(29) and y_t (red line) given by Eq.(30). (b) Wavelet coherency - the cone of influence is shown with a thick black line. Coherency ranges form blue (low coherency) to red (high coherency). (c) - (d) Phases and phase-difference. The green line represents the y_t phase, the blue line the x_t phase and the red line represents the phase-difference between y_t and x_t; (c) is for the $2.5 \sim 3.5$ year frequency band and (d) for the $5 \sim 7$ year frequency band.

the regions of strong coherency around those frequencies. That the 3-year cycles are in phase (positively correlated) is revealed by the phase-difference (red line in the upper right graph), which is consistently situated between $-\pi/2$ and $\pi/2$. Finally, we can see that the 3-year y_t cycle was leading for the first half of the time and lagging in the second half, by noting that in the first half of the sample the phase-difference is between zero and $\pi/2$, while in the second half it is between $-\pi/2$ and zero.

Looking at the 6-year cycle, we observe that the series are out of phase (negatively correlated) with x_t leading in the first half and y_t leading in the second half of the sample.

In this example, we observe that not only the wavelets are adequate to capture structural breaks and transient relations, but that they can also distinguish between different relations that occur at the same time but at distinct frequencies.

4 Higher Order Coherencies: Partial and Multiple Coherencies

When more than two series are given and when the power in one of them is to be determined or when the association between two of them is to be assessed, it is often important to account for the interaction with the other series. In this context, and in analogy with the Fourier spectral case, it makes sense to introduce the concepts of wavelet multiple coherency and wavelet partial coherency. As stated by Priestley (1992, p.681): "(...) the whole 'apparatus' of multivariate linear regression theory can be taken over (almost unchanged) and applied to the study of multivariate spectral relationships. In particular, the ideas of 'multiple correlation' and 'partial correlation' (...) have immediate analogues in the frequency domain, where they become 'multiple coherency' and 'partial coherency'."

The concepts of multiple wavelet coherency and partial wavelet coherency are simple generalizations of the corresponding concepts of (Fourier) multiple coherency and partial coherency to the time-frequency plane. The formulas used to compute these quantities are, therefore, strongly based on the corresponding formulas of multiple correlation and partial correlation.

4.1 Notations

Let p ($p > 2$) time series $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_p$, with $\boldsymbol{x}_i = \{\boldsymbol{x}_{in}, n = 0, \ldots, T-1\}$, be given. Just as in the case of ordinary wavelet coherency, to compute multiple and partial wavelet coherencies it is necessary to perform a smoothing operation on the cross-spectra. We will denote by S_{ij} the smoothed version of W_{ij}, i.e.

$$S_{ij} = S\left(W_{ij}\right), \tag{31}$$

where S is a certain smoothing operator.

Let $\mathscr{S}$ denote the $p \times p$ matrix of all the smoothed cross-wavelet spectra S_{ij}, i.e.

$$\mathscr{S} = \begin{bmatrix} S_{11} & S_{12} & \cdots & S_{1p} \\ S_{21} & S_{22} & \cdots & S_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ S_{p1} & S_{p2} & \cdots & S_{pp} \end{bmatrix} \tag{32}$$

The above matrix depends on the specific value (τ, s) at which the spectra are being computed. Also note that the matrix $\mathscr{S}$ is an Hermitian matrix, i.e. $\mathscr{S} = \mathscr{S}^H$ with the symbol H denoting conjugate transpose; in fact, we have $S_{ij} = S_{ji}^*$, for all $i \neq j$ and $S_{ii} = S\left(|W_i|^2\right)$ is a real (positive) number for all i.

Finally, for a given matrix A, we denote by A_{ij}^d the cofactor of the element in position (i, j) of A, i.e.

$$A_{ij}^d = (-1)^{(i+j)} \det A_i^j,$$

where A_i^j denotes the sub-matrix obtained from A by deleting its i^{th} row and j^{th} column. For completeness, we also use the notation $A^d := \det A$.

4.2 Multiple Wavelet Coherency

The *squared multiple wavelet coherency* between the series $\boldsymbol{x_1}$ and all the other series $\boldsymbol{x_2}, \ldots, \boldsymbol{x_p}$ will be denoted by $R^2_{1(23\ldots p)}$ and is given by the formula

$$R^2_{1(23\ldots p)} = 1 - \frac{\mathscr{S}^d}{S_{11}\mathscr{S}^d_{11}} \tag{33}$$

The *multiple wavelet coherency* $R_{1(23\ldots p)}$ is defined as the positive square root of the above quantity. We will also use the shorter notation $R^2_{1(\boldsymbol{q})}$ for $R^2_{1(2\ldots p)}$ i.e. we will simply write $\boldsymbol{q}$ to designate all the series indexes except the index 1.

4.3 Partial Wavelet Coherency

The *complex partial wavelet coherency* of $\boldsymbol{x}_1$ and $\boldsymbol{x}_j$ $(2 \leq j \leq p)$ allowing for all the other series will be denoted by $\varrho_{1\,j.\boldsymbol{q}_j}$ and is given by

$$\varrho_{1\,j.\boldsymbol{q}_j} = -\frac{\mathscr{S}^d_{j1}}{\sqrt{\mathscr{S}^d_{11}\,\mathscr{S}^d_{jj}}}, \tag{34}$$

where $\boldsymbol{q}_j$ is a short notation for all the indexes in $\boldsymbol{q}$ excluding the index j, i.e. $\boldsymbol{q}_j = \{2, \ldots, p\} \setminus \{j\}$.

The *partial wavelet coherency* of $\boldsymbol{x}_1$ and $\boldsymbol{x}_j$ allowing for all the other series, denoted by $r_{1\,j.\boldsymbol{q}_j}$, is defined as the absolute value of the above quantity, i.e.

$$r_{1\,j.\boldsymbol{q}_j} = \frac{|\mathscr{S}^d_{j1}|}{\sqrt{\mathscr{S}^d_{11}\mathscr{S}^d_{jj}}}, \tag{35}$$

whilst the *squared partial wavelet coherency* of $\boldsymbol{x}_1$ and $\boldsymbol{x}_j$ allowing for all the other series is given by

$$r^2_{1\,j.\boldsymbol{q}_j} = \frac{|\mathscr{S}^d_{j1}|^2}{\mathscr{S}^d_{11}\mathscr{S}^d_{jj}}. \tag{36}$$

4.4 Formulas in Terms of Simple (Complex) Coherencies

The above formulas for the multiple and partial coherencies were given in terms of the smoothed spectra S_{ij}. We can also define these coherencies in terms of simple complex coherencies (i.e. wavelet complex coherencies between pairs of series).

Corresponding to the matrix $\mathscr{S}$, we now consider the matrix $\mathscr{C}$ of all the complex wavelet coherencies ϱ_{ij}, i.e.

$$\mathscr{C} = \begin{bmatrix} 1 & \varrho_{12} & \cdots & \varrho_{1p} \\ \varrho_{21} & 1 & \cdots & \varrho_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \varrho_{p1} & \varrho_{p2} & \cdots & 1 \end{bmatrix} \tag{37}$$

Note that $\varrho_{jj} = \frac{S(W_{jj})}{(S(|W_j|^2)S(|W_j|^2))^{1/2}} = \frac{S(|W_j|^2)}{S(|W_j|^2)} = 1$. As $\mathscr{S}$, the matrix $\mathscr{C}$ is also an Hermitian matrix, i.e. $\varrho_{ij} = \varrho^*_{ji}$.

Then, we can define the multiple and partial wavelet coherencies by the following alternative formulas:

$$R^2_{1(\boldsymbol{q})} = 1 - \frac{\mathscr{C}^d}{\mathscr{C}^d_{11}}, \tag{38}$$

$$\varrho_{1\,j.\boldsymbol{q}_j} = -\frac{\mathscr{C}^d_{j1}}{\sqrt{\mathscr{C}^d_{11}\mathscr{C}^d_{jj}}}, \tag{39}$$

and

$$r_{1\,j.\boldsymbol{q}_j} = \frac{|\mathscr{C}^d_{j1}|}{\sqrt{\mathscr{C}^d_{11}\mathscr{C}^d_{jj}}}. \tag{40}$$

The proof of the above results is an application of the multilinear character of a determinant; we illustrate the result for the case of the squared multiple coherency in Appendix A.

4.5 Expression of Multiple Coherency in Terms of Partial Coherencies

The squared multiple coherency can be expressed in terms of squared partial coherencies, by using the following formula:

$$1 - R^2_{1(2\ldots p)} = (1 - r^2_{12})(1 - r^2_{13.2}) \ldots (1 - r^2_{1p.23\ldots(p-1)}). \tag{41}$$

The proof of the above result is given in Appendix A.

4.6 Partial Phase-Difference

Having defined the complex partial wavelet coherency $\varrho_{1\,j.\mathbf{q}_j}$ between the series $\boldsymbol{x}_1$ and the series $\boldsymbol{x}_j$, after removing the influence of all the remaining series, we now define the *partial phase-delay (phase-difference) of* $\boldsymbol{x}_1$ *over* $\boldsymbol{x}_j$*, given all the other series*, as the angle of $\varrho_{1\,j.\mathbf{q}_j}$. We will denote this phase-difference by $\phi_{1\,j.\mathbf{q}_j}$, i.e.

$$\phi_{1\,j.\boldsymbol{q}_j} = \text{Arctan}\left(\frac{\Im(\varrho_{1\,j.q_j})}{\Re(\varrho_{1\,j.q_j})}\right) \tag{42}$$

4.7 Formulas For Three Variables

Let us illustrate the use of the above formulas for the case where we just have three series $\boldsymbol{x}_1$, $\boldsymbol{x}_2$ and $\boldsymbol{x}_3$. In the appendix we show that, in this case, we have

$$R^2_{1(2\,3)} = \frac{R^2_{12} + R^2_{13} - 2\Re\left(\varrho_{12}\,\varrho_{23}\,\varrho^*_{13}\right)}{1 - R^2_{23}} \tag{43}$$

For the complex partial wavelet coherency $\varrho_{1\,2.3}$, formula (39) gives

$$\varrho_{1\,2.3} = \frac{\varrho_{12} - \varrho_{13}\varrho^*_{23}}{\sqrt{(1 - R^2_{13})(1 - R^2_{23})}} \tag{44}$$

4.8 Example 3: Partial Coherency and Phase-Difference

Consider three time series that share two common cycles, with some leads and delays.

$$x_t = \sin\left(\tfrac{2\pi}{3}t\right) + 3\sin\left(\tfrac{2\pi}{6}t\right) + \varepsilon_{x,t}$$

$$y_t = 4\sin\left(\tfrac{2\pi}{3}(t+\tfrac{5}{12})\right) + 3\sin\left(\tfrac{2\pi}{6}(t-\tfrac{10}{12})\right) + \varepsilon_{y,t}\ , \qquad t = 0, \frac{1}{12}, \frac{2}{12}, \ldots, 50.$$

$$z_t = 3\cos\left(\tfrac{2\pi}{6}t\right)$$

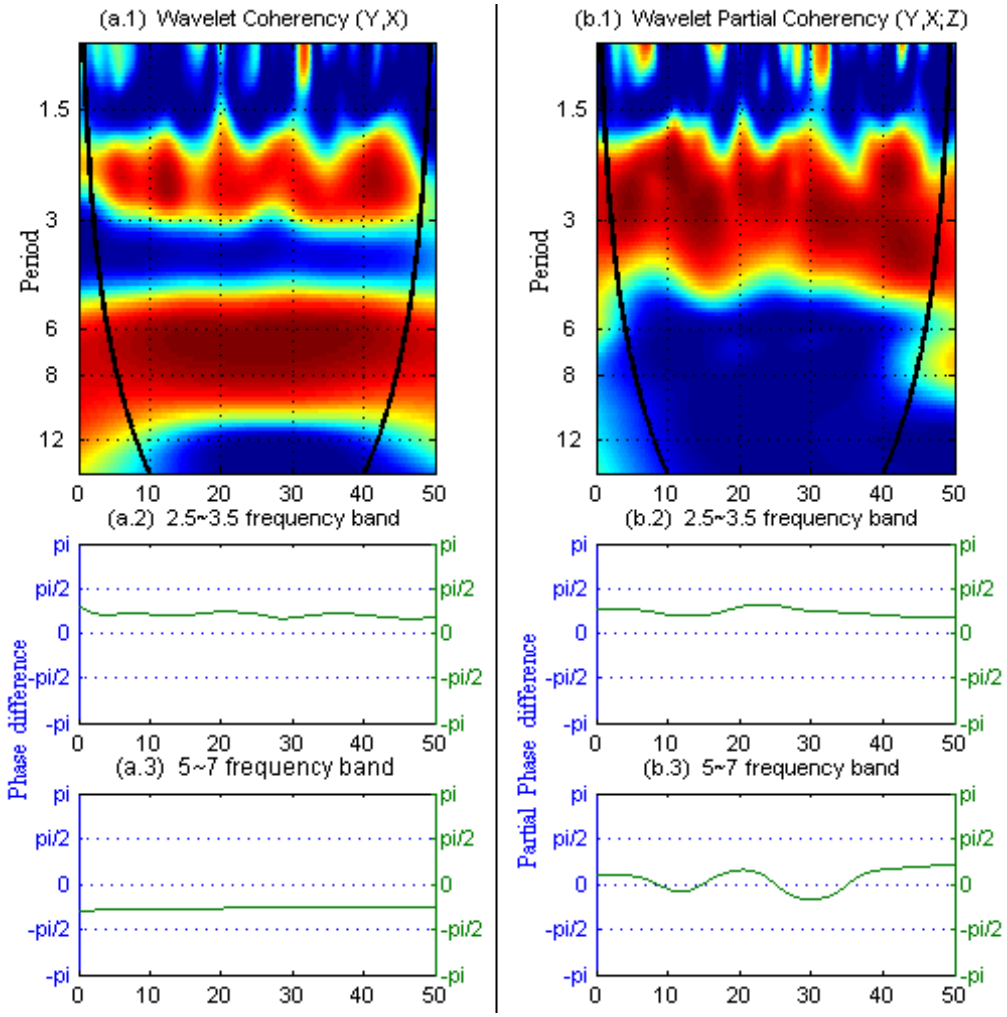

Figure 6: (a) Wavelet coherency between x_t and y_t; (b) Wavelet partial coherency between x_t and y_t, after controlling for z_t; (a.2) - (a.3) Phase-difference; (b.2) - (b.3) Partial phase-difference.

Looking at the formulas, it is clear that x_t and y_t share 3-year and 6-year cycles. While y_t leads x_t in the shorter period cycle, the opposite happens in the longer period cycle. In this example, we added a third variable, z_t, which shares the six year cycle with x_t and z_t. The cycles are positively correlated.

In Figure 6, on the left, we compute the wavelet coherency between y_t and x_t. On the right, we compute the partial wavelet coherency, after controlling for z_t. The results are what one would

expect. While the wavelet coherency and phase-difference (on the left) captures the relation between both cycles, on the right, after controlling for z_t, which shares the 6 year cycle with x_t an y_t, only the three year cycle relations are captured. And actually these relations are captured even better than with the simple coherency. This is to be expected: when there are several cycles influencing the series at the same time, after controlling for some of those, we should get a cleaner view of the others.

The concept of wavelet partial coherency is an extension of the concept of wavelet coherency just like partial correlation is an extension of the simple correlation. With this tool, one can move beyond uni and bivariate wavelet analysis to higher order variate wavelet analysis.

5 Significance Tests

As with other time series methods, it is important to assess the statistical significance of the results obtained by wavelet analysis. The seminal paper by Torrence and Compo (1998) is one of the first works to discuss significance testing for wavelet and cross-wavelet power. Based on a large number of Monte Carlo simulations, Torrence and Compo concluded that the local wavelet power spectrum of a white noise or an AR(1) process, normalized by the variance of the time series, is well approximated by a chi-squared distribution. Torrence and Compo also derived empirical distributions for cross-wavelet power. If speed is an issue, then having these distributions derived is a plus. However, if computer time is not a constraint, given that these distributions were derived by Monte Carlo simulations, then one might as well just do the Monte Carlo simulations directly. Ge (2007 and 2008) reconsiders the discussion of the significance testing for the wavelet, cross-wavelet power and wavelet coherency. The author concentrates on the use of a specific wavelet (the Morlet wavelet) and, assuming a Gaussian white noise process, analytically derive the corresponding sampling distributions. However, these sampling distributions were shown to be highly dependent on the local covariance structure of the wavelet, a fact that makes the significance levels intimately related to the specific wavelet family used, meaning that they cannot be generalized. Naturally, no work has been done on significance testing for the partial coherency, as, to the best of our knowledge, this measure has not been introduced elsewhere. Maraun, Kurths and Holschneider (2007) argued that pointwize significance tests, like the ones described, generate too many false positives. They proposed an areawise test which aims at correcting false positives of pointwise tests, based on the area and shape of the significant regions. Lachowicz (2009), however, shows that some more work needs to be done in this area.

In our examples and in our toolbox, the tests of significance are either based on very simple Monte Carlo simulations or bootstrapping. We fit an ARMA(p, q) model and then construct new samples by bootstrap or by drawing errors from a Gaussian distribution. In the first option, we use the very basic bootstrap technique described in Section 2.1 of Berkowitz and Kilian (2000).

To our knowledge, there are no good statistical tests for the phase-difference. In fact, Ge (2008) showed that, under the null of no linear relation between two variables, the phase angle will be uniformly distributed. Hence it will be dispersed between -π and π. Because of that, Ge argues that one should not use significance tests for the wavelet phase-difference. Instead, its analysis should be complemented by inspection of the coherence significance.

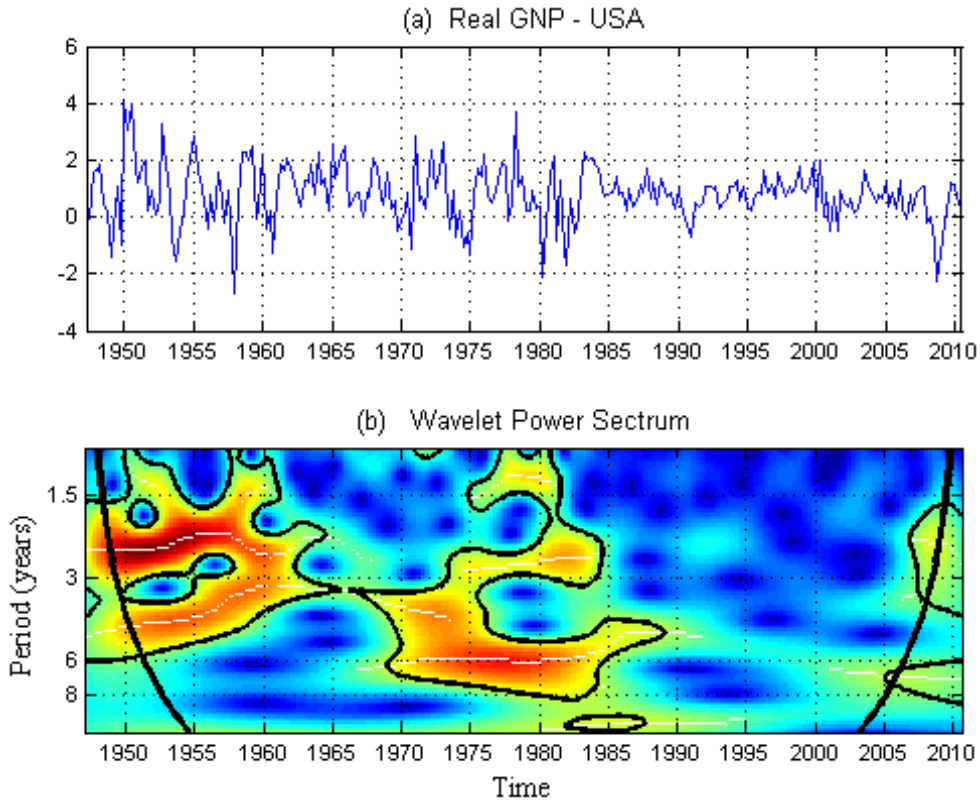

Figure 7: (a) GNP (quarterly) growth rate for the United States; (b) Wavelet power spectrum - the black contour designates the 5% significance level based on an ARMA(1,1) null. The cone of influence, which indicates the region affected by edge effects, is shown with a thick black line. The color code for power ranges from blue (low power) to red (high power). The white lines show the maxima of the undulations of the wavelet power spectrum.

6 Three Applications

6.1 The Great Moderation in the United Sates

In Figure 7 (a), we have the real GNP (quarterly) growth rate for the United States, from 1947q2 until 2010q3. In Figure 7 (b), one can observe the wavelet power spectrum. At business cycle frequencies, the wavelet power was high and statistically significant, until early 1960s. After that, the volatility at all frequencies steadily decreased, with an exception between mid 1970s and 1984, when the variance at the business cycle frequency (1.5 to 8 years) was quite high again, probably as a result of the severe oil crisis that hit the world economy in 1973 and 1979 and lasted until the early 1980s.

These results for the macroeconomic volatility in the United States are compatible with the results of Gallegati and Gallegati (2007), who studied this issue for the G-7 economies using discrete wavelet analysis. To be more precise, they use the MODWT to decompose the output time series in scales (or frequencies). Then they are able to decompose the wavelet variance on a scale-by-scale basis. Finally, they estimate each scale variance for four different decades. They also conclude that volatility in 1970s increases, probably due to the oil shocks, at every relevant scale/frequency. After that, volatility decreases. This example is appropriate to illustrate that continuous and discrete

wavelet analysis are quite complementary, as one would expect.

These results from wavelet analysis help to qualify some of the results present in the literature. The literature has identified 1984 as the year that marks the beginning of the Great Moderation (Kim and Nelson 1999; McConnell and Pérez-Quirós 2000). In reality, we can observe that this Great Moderation may have started sometime earlier. It was in the early 1960s that the volatility started to decrease. It then was revived, due to the oil shocks, at the business cycle frequency in the 1970s, however this increase was temporary. These results are in line with Blanchard and Simon (2001) who have argued that the large shocks in the 1970s and the deep contraction in early 1980s hide from view the longer term volatility decline that began a few decades before.

As one would expect, given the turbulence of the last years, after 2007 there is again evidence that volatility is increasing, suggesting that the 'Great Moderation' is not so great anymore. We see this because the wavelet power spectrum becomes statistically significant in the late 2000s at 1.5 to 5 years frequencies. Although part of this region may be affected by edge effects (because it is under the effect of the cone of influence) it is also true that a part of it is not affected by those edge effects. Finally, one should also keep in mind that, because of the zero padding, this influence will tend to underestimate, not overestimate, the power spectrum.[19]

6.2 Stock Markets: Who is the Leader?

In this application, we illustrate how one can use wavelet coherency and phase-difference analysis to study synchronism between two time series. To illustrate this, we assess the coordination between three different stock market indices. We collect monthly data on the Price Index for FTSE All-Share (United Kingdom), the S&P 500 (United States) and the DAX (Germany). This issue of the international stock markets comovement has drawn some attention in the literature. In particular, after October 1987, when several markets fell together, despite heterogenous economic conditions (King and Wadhwani 1990). In spite of this episode, King, Sentana and Wadhwani (1992) found evidence against the hypothesis of internationally integrated capital markets. More recently, Forbes and Rigobon (2002) found evidence of strong international interdependence linkages in the 1997 East Asian crises, the 1994 Mexican peso collapse, and the 1987 U.S. stock market crash and Brooks and Del Negro (2004) shows that there is a rise in comovement across national stock markets after the mid-1990s among the major developed countries. For more on this, the reader can see Rua and Nunes (2009)

In Figure 8, we can see the wavelet coherency between S&P and FTSE and between S&P and DAX.[20] It is apparent that New York shows more regions of high coherency with the London stock market than with Frankfurt. This is particularly evident when one focus on the decades of 1980 and 1990 at the higher frequency band (1∼4 years). These results are in line with Rua and Nunes (2008) who also concluded that the US and UK stock markets have a high degree of comovement over the last forty years.

These pictures suggest that the UK and the US stock markets became more synchronized in 1985, synchronization that was extended to Germany only in the decade of 1990. An interesting conclusion, not present in Rua and Nunes (2008), arises when one looks at the phase-differences.

[19] In Appendix A we explain the reason for these border distortions and explain how we deal with it.

[20] The coherency between FTSE and Dax, not shown, is very similar to the coherency between S&P and DAX.

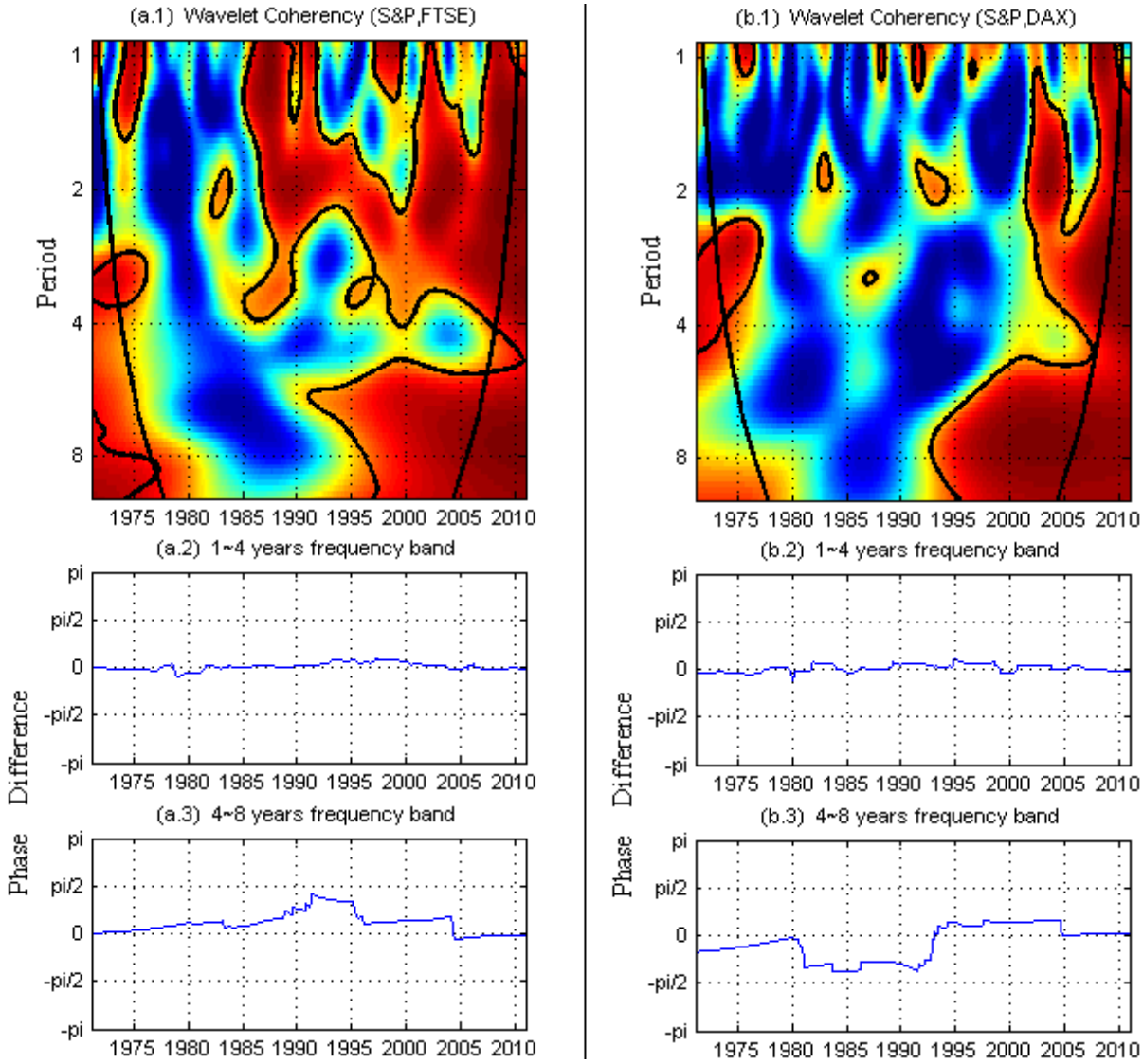

Figure 8: (a.1) - (b.1) Wavelet coherency between S&P and FTSE and between S&P and DAX; the black contour designates the 5% significance level based on an ARMA(1,1) null; (a.2) - (b.2) Phase-differences at 1∼4 years frequency band; (a.3) - (b.3) Phase-differences at 4∼8 years frequency band.

While, in the shorter run frequencies, phases are very much aligned and, therefore, the phase-difference is very close to zero (more often positive than negative, one might add), when one looks at 4 to 8 year period frequencies, the phase-difference is consistently above zero, meaning that the US stock market leads the other stock markets.[21] The same analysis, not shown, comparing the UK and the German stock markets would lead to the conclusion that these two markets move very much together, with no noticeable lead or delay in the regions of high coherency.

[21] It is true that until early 1990s the phase-difference between S&P and DAX is negative, however this is a period of low coherency and, therefore, there is not much meaning attached to the phase-difference.

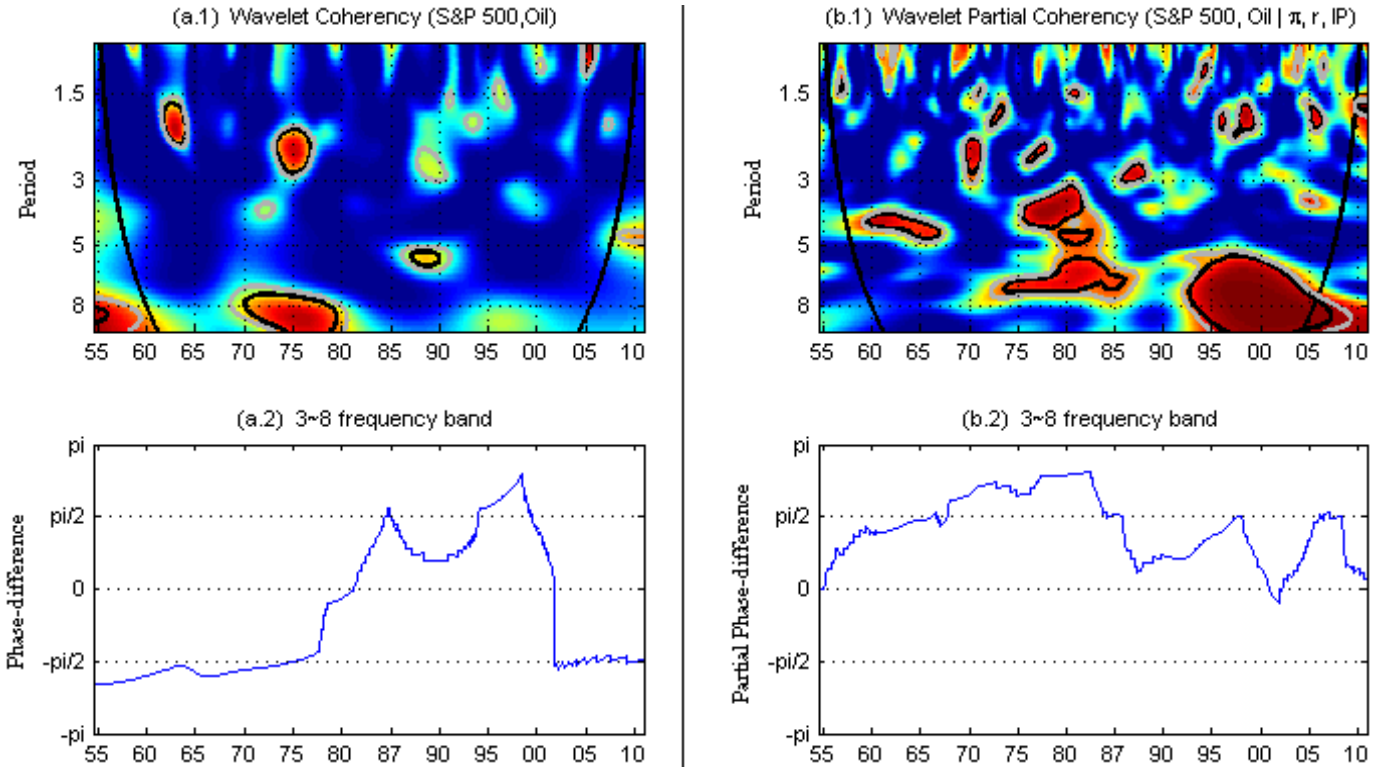

Figure 9: (a.1) Wavelet coherency between S&P and oil prices; (b.1) Wavelet partial coherency between S&P and oil prices, after controlling for inflation, real interest rates and industrial production. The black (grey) contour designates the 5% (10%) significance level based on an ARMA(1,1) null; (a.2) Phase-difference at $3 \sim 8$ years frequency band; (b.2) Partial phase-differences at $3 \sim 8$ years frequency band.

6.3 Stock Markets and Oil Prices.

The macroeconomic impact of oil price shocks is the subject of innumerous papers and modelling its effects is not as trivial as one may think (see Aguiar-Conraria and Wen 2007 and Kilian 2008). Aguiar-Conraria and Soares (2011a) have already applied wavelet analysis to study the relation between oil prices and the macroeconomy. There is also a smaller literature devoted to the impact of oil shocks in the stock markets. Sadorsky (1999) and Ciner (2001), for example, found that increases in oil prices had, in general, negative impacts on stock market returns. In a different direction were the conclusions of Huang, Masulis and Stoll (1996). They concluded that linkages between oil shocks and the financial markets were, at best, weak. Kilian and Park (2009) reach more subtle conclusions. They basically conclude that if oil price increases are the results of oil supply shocks (or expectation that there will be a supply shortage) then their impact on the stock market is negative. However, an increase in global aggregate demand will result in both higher real oil prices and higher stock prices. We now use higher order wavelet tools to briefly study the linkages between oil prices and stock market returns. It should become clear that moving from bivariate wavelet analysis to higher order analysis will avoid reaching erroneous conclusions.

We gathered monthly data, running from July 1954 to December 2010, on several variables: S&P-500 Stock Monthly Returns, (log) Oil Prices[22], (log) Industrial Production, CPI inflation and

[22] For the oil price, we considered a nonlinear transformation *a la* Hamilton (2003). For each period, we consider the

the Effective Federal Funds Real Rate. In Figure 9, on the left, we have the wavelet coherency and phase-difference between the stock market returns and oil prices; on the right, we have the partial wavelet coherency and phase-difference after controlling for the other variables.

It is difficult to make sense of the picture on the left. Regions of high coherency are very scarce and the phase-difference is not meaningful in these situations. If anything, one would conclude that there are no relevant linkages between oil prices and the stock market. However, once the other variables are controlled for, the picture becomes clearer. On the right we can see that there regions of high coherency between the mid-1970s and mid-1980s along the $3 \sim 8$ year period frequency band and again, at lower frequencies after the early 1990s. The phase-difference reveals some interesting insights. In the 1970s and 1980s (at least until the market crash of 1987), the partial phase-difference is between $\pi/2$ and π, implying an anti-phase relation with the oil prices leading. This suggests that the oil price increases precede stock market downturns. This is totally compatible with the results of Kilian and Park (2009), given that in the decade of 1970 and early 1980 the oil crises were clearly on the supply side. There is a structural change after that. In the 1990s and 2000s the regions of high coherency are situated at lower frequencies and, at the same time, the phase-difference is consistently between 0 and $\pi/2$, suggesting that the series are now in phase, with the stock market leading. This result reinforces the demand side hypothesis. An increase in economic activity will, naturally, be reflected in the stock market and pressure on oil demand will lead to oil price increases, lending further support to the conclusions of Kilian (2009), Kilian and Park (2009) and Baumeister and Peersman (2008).

7 Conclusion

In this paper, we argued that continuous wavelet analysis can be a very useful tool to analyze business cycles. We had three main objectives: (1) to present a self-contained summary on the most relevant theoretical results related to the continuous wavelet transform, (2) to describe how such transforms can be implemented in practice, and (3) to generalize the concept of multiple and partial coherencies to the time-frequency framework provided by wavelet analysis. We also presented some results on a new family of wavelets, the generalized Morse wavelets, which are becoming more popular in other scientific fields and allow for more flexibility than the popular Morlet wavelet, while keeping some of its nice properties.

To illustrate the potentialities of wavelet analysis and to provide some easy to do examples, we worked out three constructed examples and three real data applications. The constructed examples were put together to show (1) how wavelet analysis can easily capture transient cycles that are not stable across time and frequencies and to uncover information that would be difficult to extract with the more traditional Fourier analysis; (2) how cross wavelets can capture transient relationships between two time series at different frequencies; and (3) to illustrate the usefulness of a new concept: higher order coherencies.

We also provide three applications with real data. In one of them, we applied the wavelet power

maximum price achieved in the last three years. As Hamilton argues, by doing this any oil price increase is captured while only permanent oil price decreases are considered. This an efficient way to capture the asymmetric effects of oil price changes.

spectrum to study the post-war business cycle volatility in the United States, by looking at real GNP growth rates. We concluded that while it is true that business cycles were very active during the 70s and early 80s (after the oil price shocks in the 70s), it is also true that our results support the view of Blanchard and Simon (2001), according to whom the large shocks in the 1970s disguised the fact that the moderation had begun a few decades earlier. In our second application, with the help of cross wavelet analysis, we studied synchronism in international stock market returns. We showed that comovement between the US and the UK markets has become stronger after 1985, while the comovement between these stock markets and the German stock market has only increased after mid-1990s. We have also shown that at low frequencies ($4 \sim 8$ year period frequencies) the US stock market leads the other mentioned stock markets.

Finally, on our third empirical application, we use partial wavelet coherencies and partial phase-difference to study the linkages between oil prices and the stock markets. It is clear that some of the conclusions we reached would be very difficult to spot had we just used bivariate cross wavelet analysis.

Attached to this paper, there is a Matlab toolbox implementing the referred wavelet tools, which the researcher can freely use and adapt to his/her own research.

8 Appendix A

8.1 Discrete Computations

Suppose we sample the series $x(t)$ with a fine enough sample interval δt to avoid aliasing (i.e. assume that $X(\omega) \approx 0$ for $|\omega| > \frac{2\pi}{2\delta t} = \frac{\pi}{\delta t}$) and use the shorthand notation $x_n = x(n\delta t); n = 0, \ldots, N-1$. Also, let $\boldsymbol{x} = \{x_n; n = 0, \ldots, N-1\}$. With N even, formula (6) can be discretized as

$$\begin{aligned} W_x(\tau, s) &\approx \frac{\sqrt{s}}{2\pi} \int_{-\frac{\pi}{\delta t}}^{\frac{\pi}{\delta t}} X(\omega) \Psi^*(s\omega)\, e^{i\omega\tau} d\omega \\ &\approx \frac{\sqrt{s}}{2\pi} \frac{2\pi}{N\delta t} \sum_{k=-(N/2)+1}^{N/2} X\left(\frac{2\pi k}{N\delta t}\right) \Psi^*\left(\frac{s2\pi k}{N\delta t}\right) e^{i\frac{2\pi k}{N\delta t}\tau} \\ &= \frac{\sqrt{s}}{N\delta t} \sum_{k=-(N/2)+1}^{N/2} X\left(\frac{2\pi k}{N\delta t}\right) \Psi^*\left(\frac{s2\pi k}{N\delta t}\right) e^{i\frac{2\pi k}{N\delta t}\tau} \end{aligned}$$

But, $X\left(\frac{2\pi k}{N\delta t}\right) \approx \delta t\, \hat{x}_k$, where $\hat{x}_k = \sum_{n=0}^{N-1} x_n e^{-i2\pi nk/N}$ is the kth element of the Discrete Fourier Transform (DFT) of the N-vector $(x_0, \ldots, x_{N-1})$; see e.g. Brémaud (2002). Hence, we obtain a discretized form of the CWT of the discrete time series $\boldsymbol{x} = \{x_n : 0, \ldots, N-1\}$:

$$\begin{aligned} W_{\boldsymbol{x}}(\tau, s) &= \frac{\sqrt{s}}{N} \sum_{k=0}^{N/2} \hat{x}_k \Psi^*(s\frac{2\pi k}{N\delta t}) e^{\frac{i2\pi k\tau}{N\delta t}} \\ &\quad + \frac{\sqrt{s}}{N} \sum_{k=(N/2)+1}^{N-1} \hat{x}_k \Psi^*(s\frac{2\pi(k-N)}{N\delta t}) e^{\frac{i2\pi(k-N)\tau}{N\delta t}} \end{aligned}$$

where we used the periodicity $\hat{x}_k = \hat{x}_{k-N}$. When $\tau = m\,\delta t;\ m = 0, \ldots, N-1$, we get

$$\begin{aligned} W_{\boldsymbol{x}}(m\delta t, s) &= \frac{\sqrt{s}}{N} \sum_{k=0}^{N/2} \hat{x}_k \Psi^*(s\frac{2\pi k}{N\delta t}) e^{\frac{i2\pi km}{N}} \\ &\quad + \frac{\sqrt{s}}{N} \sum_{k=(N/2)+1}^{N-1} \hat{x}_k \Psi^*(s\frac{2\pi(k-N)}{N\delta t}) e^{\frac{i2\pi km}{N}} \\ &= \frac{\sqrt{s}}{N} \sum_{k=0}^{N-1} \hat{x}_k \Psi^*(s w_k) e^{\frac{i2\pi km}{N}} \end{aligned}$$

where

$$w_k = \begin{cases} \frac{2\pi k}{N\delta t}, & k = 0, 1, \ldots, \frac{N}{2}, \\ \frac{2\pi(k-N)}{N\delta t}, & k = \frac{N}{2}+1, \ldots, N-1. \end{cases} \tag{45}$$

In practice, naturally, the wavelet transform is computed only for a selected set of scale values $s \in \{s_\ell, \ell = 0, \ldots, F-1\}$ (corresponding to a certain choice of frequencies ω_ℓ). Hence, our computed wavelet spectrum of the discrete-time series $\boldsymbol{x}$ will simply be a $F \times N$ matrix $\mathbf{W}_{\boldsymbol{x}}$ (wavelet spectral

matrix) whose (ℓ, m) element is given by

$$\mathbf{W}_{\boldsymbol{x}}(\ell, m) = \frac{\sqrt{s_\ell}}{N} \sum_{k=0}^{N-1} \hat{x}_k \Psi^*(s_\ell w_k) e^{\frac{i2\pi km}{N}}, \tag{46}$$

with w_k given by (45).[23] The above formula is an efficient formula for computing the CWT, since, for each scale s_ℓ, its right-hand side is simply the inverse DFT of the sequence $(z_0^\ell, \ldots, z_{N-1}^\ell)$ where

$$z_k^\ell := \sqrt{s_\ell}\, \hat{x}_k\, \Psi^*(s_\ell\, w_k); k = 0, \ldots, N-1,$$

and can, therefore, be calculated using an inverse FFT.

Naturally, all the formulas given previously for other wavelet measures, such as the cross-wavelet transform, the wavelet coherency and the phase-difference have discrete counterparts. For example, corresponding to formula (23) for the cross-wavelet transform, we have a discretized version $\mathbf{W}_{\boldsymbol{xy}}(\ell, m) = \mathbf{W}_{\boldsymbol{x}}(\ell, m)\mathbf{W}_{\boldsymbol{y}}^*(\ell, m)$.Formula (27) for the phase-difference is discretized as $\phi_{\boldsymbol{xy}}(\ell, m) = \text{Arctan}\left(\frac{\Im(\mathbf{W}_{\boldsymbol{xy}}(\ell,m))}{\Re(\mathbf{W}_{\boldsymbol{xy}}(\ell,m))}\right)$.

8.1.1 Choice of scales

The scales are usually chosen as fractional powers of 2:

$$s_\ell = s_0\, 2^{\frac{\ell}{n_V}}; \ell = 0, 1, \ldots, n_V \times n_O, \tag{47}$$

where n_O denotes the number of octaves (i.e. powers of two) and n_V the number of voices calculated per octave (i.e., $F = n_V \times n_O + 1$).

8.1.2 Cone of influence

As with other types of transforms, the CWT applied to a finite length time-series inevitably suffers from border distortions; this is due to the fact that the values of the transform at the beginning and the end of the time-series are always incorrectly computed, in the sense that they involve missing values of the series which are then artificially prescribed. When using the formula (46), a periodization of the data is assumed. However, before implementing formula (46), one usually pads the series with zeros, to avoid wrapping. Since the "effective support" of the wavelet at scale s is proportional to s, these edge-effects also increase with s. The region in which the transform suffers from these edge effects is called the cone of influence (COI). In this area of the time-frequency plane the results are subject to border distortions and have to be interpreted carefully.

8.2 Higher order cross-wavelets: some results

8.2.1 Proof that $R^2_{1(\boldsymbol{q})} = 1 - \frac{\mathscr{C}^d}{\mathscr{C}^d_{11}}$

Since $S_{ii} = S\left(W_i W_i^*\right) = S\left(|W_i|^2\right)$ is a non-negative quantity, let $s_i := \sqrt{S_{ii}}$. Recalling the formula which expresses the complex coherencies in terms of the wavelet spectra (25) and the notation of

[23] We choose to make the row indexes of the matrix correspond to scales and the columns to times, so that the plots of this matrix will naturally lead to times in the x-axis and frequencies (or periods) in the y-axis.

equation (31), we see that $\varrho_{ij} = \frac{S_{ij}}{s_i s_j}$. Hence,

$$\mathscr{C}^d = \det \mathscr{C} = \det \begin{bmatrix} 1 & \frac{S_{12}}{s_1 s_2} & \cdots & \frac{S_{1p}}{s_1 s_p} \\ \frac{S_{21}}{s_2 s_1} & 1 & \cdots & \frac{S_{2p}}{s_2 s_p} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{S_{p1}}{s_p s_1} & \frac{S_{p2}}{s_p s_2} & \cdots & 1 \end{bmatrix}$$

By using a well-known property of determinants (if a column or a row of a matrix is multiplied by a constant, the determinant is multiplied by that constant), we have that

$$\mathscr{C}^d = \frac{1}{s_1^2 s_2^2 \dots s_p^2} \mathscr{S}^d.$$

Also, since $\mathscr{C}_{11}^d$ is the determinant of the matrix obtained by deleting the first row and first column of $\mathscr{C}$, it is immediate to conclude that

$$\mathscr{C}_{11}^d = \frac{1}{s_2^2 \dots s_p^2} \mathscr{S}_{11}^d.$$

Hence, we have

$$1 - \frac{\mathscr{C}^d}{\mathscr{C}_{11}^d} = 1 - \frac{s_2^2 \dots s_p^2}{s_1^2 s_2^2 \dots s_p^2} \frac{\mathscr{S}^d}{\mathscr{S}_{11}^d} = 1 - \frac{\mathscr{S}^d}{\sigma_1^2 \mathscr{S}_{11}^d} = 1 - \frac{\mathscr{S}^d}{S_{11} \mathscr{S}_{11}^d}$$

which is precisely formula (33). The proofs of equations (39) and (40) involve similar steps.

8.2.2 Proof that $1 - R^2_{1(2\dots p)} = (1 - r^2_{12})(1 - r^2_{13.2}) \dots (1 - r^2_{1p.23\dots(p-1)})$.

Let $Q = \mathscr{C}_2^2$ (i.e. the submatrix of $\mathscr{C}$ obtained by suppressing its 2nd row and 2nd column). Then, from (33), we have

$$1 - R^2_{1(3\dots p)} = \frac{\mathcal{Q}^d}{\mathcal{Q}_{11}^d}.$$

Since,

$$1 - R^2_{1(2\dots p)} = \frac{\mathscr{C}^d}{\mathscr{C}_{11}^d},$$

we obtain

$$\frac{1 - R^2_{1(2\dots p)}}{1 - R^2_{1(3\dots p)}} = \frac{\mathcal{Q}_{11}^d}{\mathscr{C}_{11}^d} \frac{\mathscr{C}}{\mathcal{Q}^d}.$$

But, from the the definition of Q, we have that $Q^d = \mathscr{C}_{22}^d$. Moreover, by using the Jacobi's generalized theorem on determinants, we also have that

$$\mathcal{Q}_{11}^d \mathscr{C}^d = \begin{vmatrix} \mathscr{C}_{11}^d & \mathscr{C}_{12}^d \\ \mathscr{C}_{21}^d & \mathscr{C}_{22}^d \end{vmatrix},$$

since Q^d_{11} is the complementary minor of the minor $\begin{vmatrix} 1 & \varrho_{12} \\ \varrho_{21} & 1 \end{vmatrix}$ in $\mathscr{C}^d$. Hence, we get

$$\frac{1-R^2_{1(2\ldots p)}}{1-R^2_{1(3\ldots p)}} = \frac{\begin{vmatrix} \mathscr{C}^d_{11} & \mathscr{C}^d_{12} \\ \mathscr{C}^d_{21} & \mathscr{C}^d_{22} \end{vmatrix}}{\mathscr{C}^d_{11}\mathscr{C}^d_{22}} = \frac{\mathscr{C}^d_{11}\mathscr{C}^d_{22} - \mathscr{C}^d_{12}\mathscr{C}^d_{21}}{\mathscr{C}^d_{11}\mathscr{C}^d_{22}} = 1 - \frac{\mathscr{C}^d_{12}\mathscr{C}^d_{21}}{\mathscr{C}^d_{11}\mathscr{C}^d_{22}}$$

or, by using (34),

$$\frac{1-R^2_{1(2\ldots p)}}{1-R^2_{1(3\ldots p)}} = 1 - \varrho_{1\,2.\boldsymbol{q}_2}\varrho_{2\,1.\boldsymbol{q}_2}$$

But, by using the fact that $\mathscr{C}$ is Hermitian, we easily conclude that $\varrho_{2\,1.\boldsymbol{q}_j} = \varrho^*_{1\,2.\boldsymbol{q}_2}$, hence $\varrho_{1\,2.\boldsymbol{q}_2}\varrho_{2\,1.\boldsymbol{q}_2} = \varrho_{1\,2.\boldsymbol{q}_2}\varrho^*_{1\,2.\boldsymbol{q}_2} = |\varrho_{1\,2.\boldsymbol{q}_2}|^2 = r^2_{12.\boldsymbol{q}_2} = r^2_{1\,2.(3\ldots p)}$. So, we finally obtain

$$1-R^2_{1(2\ldots p)} = (1-R^2_{1(3\ldots p)})(1-r^2_{1\,2.(3\ldots p)}).$$

If we follow the same process for $1-R^2_{1(3\ldots p)}$ and replace the corresponding expression in the formula above, we obtain

$$1-R^2_{1(2\ldots p)} = (1-R^2_{1(4\ldots p)})(1-r^2_{1\,3.(4\ldots p)})(1-r^2_{1\,2.(3\ldots p)})$$

By repeating this process successively, we end up with

$$1-R^2_{1(2\ldots p)} = (1-R^2_{1p})(1-r^2_{1(p-1).p})(1-r^2_{1(p-2).((p-1)p)}\ldots(1-r^2_{12.(3\ldots p)}).$$

In the left hand side of above formula, the order in which the secondary subscripts (i.e. all the subscripts except 1) in $R^2_{1(2\ldots p)}$ are given is irrelevant; by permuting them appropriately, we can write

$$1-R^2_{1(2\ldots p)} = (1-R^2_{12})(1-r^2_{1\,3.2})(1-r^2_{1\,4.2\,3})\ldots(1-r^2_{1p.2\,3\ldots(p-1)}).$$

which is precisely formula (41), using the obvious fact that $R^2_{12} = r^2_{12}$

8.2.3 Formulas for three variables

Suppose that we just have three series $\boldsymbol{x}_1$, $\boldsymbol{x}_2$ and $\boldsymbol{x}_3$. In this case, the matrix $\mathscr{C}$ of coherencies is simply

$$\mathscr{C} = \begin{pmatrix} 1 & \varrho_{12} & \varrho_{13} \\ \varrho_{21} & 1 & \varrho_{23} \\ \varrho_{31} & \varrho_{32} & 1 \end{pmatrix}$$

Recall that this is an Hermitian matrix, and so we have: $\varrho_{ji} = \varrho^*_{ij}$ and $\varrho_{ij}\,\varrho_{ji} = \varrho_{ij}\,\varrho^*_{ij} = |\varrho_{ij}|^2 = R^2_{ij}$. Hence, we have:

$$\mathscr{C}^d_{11} = \begin{vmatrix} 1 & \varrho_{23} \\ \varrho_{32} & 1 \end{vmatrix} = 1 - \varrho_{23}\,\varrho_{32} = 1 - R^2_{23}.$$

$$
\begin{aligned}
\mathscr{C}^d &= \begin{vmatrix} 1 & \varrho_{23} \\ \varrho_{32} & 1 \end{vmatrix} - \varrho_{12} \begin{vmatrix} \varrho_{21} & \varrho_{23} \\ \varrho_{31} & 1 \end{vmatrix} + \varrho_{13} \begin{vmatrix} \varrho_{21} & 1 \\ \varrho_{31} & \varrho_{32} \end{vmatrix} \\
&= 1 - R^2_{23} - \varrho_{12}\left(\varrho_{21} - \varrho_{23}\varrho_{31}\right) + \varrho_{13}\left(\varrho_{21}\varrho_{32} - \varrho_{31}\right) \\
&= 1 - R^2_{23} - r^2_{12} - r^2_{13} + \varrho_{12}\,\varrho_{23}\,\varrho_{31} + \varrho_{21}\,\varrho_{32}\,\varrho_{13}\varrho_{32} \\
&= 1 - R^2_{23} - R^2_{12} - r^2_{13} + 2\Re\left(\varrho_{12}\,\varrho_{23}\,\varrho_{31}\right) \\
&= 1 - R^2_{23} - R^2_{12} - r^2_{13} + 2\Re\left(\varrho_{12}\,\varrho_{23}\,\varrho^*_{13}\right)
\end{aligned}
$$

Therefore, by using formula (38) for the squared multiple coherency, we get

$$
\begin{aligned}
R^2_{1(2\,3)} &= 1 - \frac{1 - R^2_{23} - R^2_{12} - R^2_{13} + 2\Re\left(\varrho_{12}\,\varrho_{23}\,\varrho^*_{13}\right)}{1 - R^2_{23}} \\
&= \frac{R^2_{12} + R^2_{13} - 2\Re\left(\varrho_{12}\,\varrho_{23}\,\varrho^*_{13}\right)}{1 - R^2_{23}}
\end{aligned}
\tag{48}
$$

For the complex partial wavelet coherency $\varrho_{1\,2.3}$, formula (39) gives

$$
\varrho_{1\,2.3} = -\frac{\mathscr{C}^d_{21}}{\sqrt{\mathscr{C}^d_{11}\,\mathscr{C}^d_{22}}}
$$

But:

$$
\mathscr{C}^d_{11} = 1 - R^2_{23} \quad \text{and} \quad \mathscr{C}^d_{22} = 1 - R^2_{13}
$$

$$
\begin{aligned}
\mathscr{C}^d_{21} &= (-1)^{(1+2)} \begin{vmatrix} \varrho_{12} & \varrho_{13} \\ \varrho_{32} & 1 \end{vmatrix} \\
&= -\left(\varrho_{12} - \varrho_{13}\varrho_{32}\right) \\
&= \left(\varrho_{12} - \varrho_{13}\varrho^*_{23}\right)
\end{aligned}
$$

Hence

$$
\varrho_{1\,2.3} = \frac{\varrho_{12} - \varrho_{13}\varrho^*_{23}}{\sqrt{(1 - R^2_{13})(1 - R^2_{23})}}
$$

9 Appendix B: Toolbox

Due to its increasing popularity and applicability into a wide range of fields, the amount of wavelet-related software has been growing at an enormous rate. Apart from commercial scientific computing software such as *Matlab*, which now integrates wavelet analysis packages,[24] there are also many free-ware wavelet software toolkits available. Naturally, these free-ware packages usually reflect the main field of interest of their authors and are oriented to use in specific applications. Lists of wavelet software, with a brief description of its main features, can be found in the following sites:

[24] Math Work's *Wavelet Toolbox* for Matlab, which includes a large collection of M-files and also GUI-based tools for wavelets. The choice of wavelets to be used in the CWT is large and the GUI interface makes the use of the toolbox specially simple. The ability to compute the wavelet coherence, cross spectrum and the cone-of-influence was only very recently added to the toolbox (Wavelet Toolbox 4.6, released in September 2010) and this was one of the reasons why the toolbox could not be used before in our work. However this toolbox still does not include significance testing, which is a major shortcoming.

– http://www.amara.com/current/wavesoft.html, and

– http://www.wavelet.org/phpBB2/gallery.php?c=Software.

There are some Matlab toolboxes that are closely related with our toolbox (i.e. which are concerned with the computation of continuous wavelet tools).

- **Torrence and Compo** – A toolbox developed by C. Torrence and G. P. Compo and available at http://atoc.colorado.edu/research/wavelets. This toolbox was developed in connection with the seminal paper by Torrence and Compo (1998) and can be considered as the "mother" of all the toolboxes developed later (all of which, including our own, incorporate or adapt some of its functions). It computes only the CWT and corresponding levels of significance. [25] The wavelets which can be used are: the Morlet wavelet, the Paul wavelet (a particular case of the GMWs) and the DOG (Derivative of Gauss) wavelet (this is a real wavelet).

- **SOWAS** – software for analysis and synthesis, developed by D. Maraun and J. Laehnemann and available at: http://tocsy.agnld.uni-potsdam.de/wavelets. It computes the CWT, XWT, wavelet coherence and phase-difference. It makes use of the wavelet software toolbox of Torrence and Compo and so, the available wavelets are the same as in Torrence and Compo's toolbox: Morlet, Paul and DOG. Smoothing in the computation of coherence is done with a box window function (both in time and in frequency). Its important contribution is the possibility of using not only pointwise significance tests (as in our case), but also areawise significance tests; see Maraun and Kurths (2004) and Maraun, Kurths and Holschneider (2007) for details.

- **Wavelet coherence package** – toolbox developed by A. Grinsted, J. C. Moore and S. Jevrejeva and available at: http://www.pol.ac.uk./home/research/waveletcoherence. It computes the CWT, XWT, wavelet coherence and phase-difference. The available wavelets are: Morlet, Paul and DOG. The smoothing used for the computation of coherence is designed for the Morlet wavelet, as described in the appendix of Torrence and Webster (1999) (and inadequate for other wavelets).

- **Wavelets_BCMC** – toolbox developed by B. Cazelles and M. Chavez.. Although the site referred by the authors in the paper by Cazelles *et. al.* (2007) is not available, the package can be obtained easily by request to one of the authors. It computes the CWT, XWT, coherence and phase-difference. The available wavelets are: Morlet, Paul and DOG. Smoothing for computation of coherence can be done with various types of window functions. This package has two main advantages. First, the statistical tests can be done by several different approaches, including using Markov hidden processes, Fourier phase randomization, theoretical white and red noise distributions, etc. Second, the output options are very rich. E.g. one can ask for the instantaneous relative variance of different frequency bands, meaning that, in each point in time, one can compute which frequencies contribute more to the overall variance.

- **ASToolbox** – this is our toolbox. We describe it next with some detail. We built on the toolboxes of Torrence and Compo and Cazelles *et al.*. This toolbox was written with economics

[25] A function to compute the wavelet coherency is available in a IDL version of this toolbox; there is also a FORTRAN version of this package.

and other social science applications in mind. The significance tests are performed using the familiar ARMA processes as the null, and can be done by simple bootstrapping or simple Monte Carlo simulations. Given that we only use analytical wavelets, simple reconstruction formulas are available, which will easily allow the proficient user to adapt the toolbox to construct wavelet based bandpass filters (for us, this is still work in progress). Our toolbox allows for the use of the generalized Morse wavelet, which encompass the most popular analytical wavelets (such as the Paul wavelet), and the Morlet wavelet. Of course, the most important contribution of our toolbox is the possibility of computing multiple and partial wavelet coherencies as well as what we called partial phase-difference.

The ASToolbox is available at http://sites.google.com/site/aguiarconraria/joanasoares-wavelets. The folder ASToolbox contains a series of Matlab functions implementing the continuous wavelet tools described in this paper. Our main objective was to collect into one single directory all the functions necessary to use these tools and also to provide some scripts illustrating their use. This, we hope, will encourage newcomers to the field to make tests with their own data and might contribute to the dissemination of the use of wavelets, not only in economics and finance, but possibly in other areas.

The folder **ASToolbox** is divided into two sub-folders:

1. **Functions** – containing all the Matlab functions. This has two sub-folders:
   - **Auxiliary** – containing some auxiliary functions to, e.g. generate surrogate series or compute Fourier spectra; it also contains a function to compute measures associated with generalized Morse wavelets.
   - **WaveletTransforms** – containing functions to compute the (analytic) wavelet transform, cross-wavelet transform, wavelet coherency, wavelet phase-difference and time-lag, multiple coherency, partial coherency and partial phase-difference.
2. **Examples** – containing Matlab scripts to generate the pictures associated with each example and application of this paper.

Some of our functions are based on (parts of) functions written by Christopher Torrence and Gilbert P. Compo (http://paos.colorado.edu/research/wavelets/) and also on some modified versions of functions written by Bernard Cazelles and Mario Chavez; Cazelles et al. (2007).

9.1 Implementation details

When implementing the transforms, some choices have, naturally, to be made. We now give a brief description of the options made in our programs. These can be modified, with very little effort, by any user.

- **Normalization**

The wavelets are normalized to have unit energy, i.e. we use $K_{\beta,\gamma} = 2\left(\frac{e\,\gamma}{\beta}\right)^{\beta/\gamma}$ for the normalizing constant, in the case of a GMW — see Lilly and Olhede (2009) — and formula (19), in the case of the Morlet wavelet.

- **Fourier factor**

The conversion of scales to frequencies is based on the energy frequency ω_ψ^E, given by (11), i.e. we use formula $\omega(s) = \frac{\omega_\psi^E}{s}$ to convert scales to angular frequencies. This, in turn, means that our Fourier factor, used to convert scales to Fourier periods, is given by

$$\mathrm{Ff} = \frac{2\pi}{\omega_\psi^E}. \tag{49}$$

- **Formula for CWT**

The implementation of the CWT (and, hence, also the XWT and the WCO) is based on the use of formula (46), together with and inverse FFT.

- **Scales**

The scales used in the CWT (XWT, WCO) are chosen as fractional powers of 2, i.e. they are of the type given by (47).

- **COI**

When implementing our algorithms, we take as decaying time to define the COI, the quantity given by the radius of the wavelet (at each scale s_ℓ), i.e. we consider

$$t_\ell = s_\ell \sigma_t.$$

But

$$t_\ell = s_\ell \sigma_t \iff t_\ell = \frac{\lambda_\ell}{\mathrm{Ff}} \sigma_t \iff \lambda_\ell = \frac{\mathrm{Ff}}{\sigma_t} t_\ell,$$

where λ_ℓ denotes the Fourier period corresponding to scale s_ℓ and Ff is the Fourier factor given by (49).

- **Smoothing**

The smoothing process involved in the coherency computations is done by convolution with window functions in time and in frequency. The type and size of the window can be selected by the user. Possible windows are: rectangular (box), triangular, Hamming, Hanning, Blackman and Bartlett.

- **Significance tests**

The tests of significance are always based on Monte Carlo simulations. The simulations use two different types of methods to construct surrogate series: (1) fitting an ARMA(p, q) model and building new samples by bootstrap or (2) fitting an ARMA(p, q) model and construct new samples by drawing errors from a Gaussian distribution. In the first option, we use the very basic bootstrap technique described in Section 2.1 of Berkowitz and Kilian (2000) and Chatterjee (1986). In the second option, the surrogates are constructed using the function 'garchsim' (univariate GARCH

process simulation) of the Econometrics Toolbox included in Matlab 2009. To fit the ARMA(p, q) to the data, we make use of the function 'garchfit' of the same toolbox. The user that does not have the Econometrics toolbox can perform significance tests by choosing an ARMA$(p, 0)$ model with bootstrap. In this case, the AR(p) model is estimated by OLS and the code is self-contained and autonomous from the Econometrics toolbox.

9.2 Software requirements

Our programs were written in Matlab 2009.b. However we were careful in writing it in such a way that it is fully compatible with version 7. Some of our programs make use of functions from the Matlab toolboxes Econometrics Toolbox and Signal Processing Toolbox. This is always explicitly stated in the function and may, in some cases, be very simply replaced by functions written by the user. If one performs significance tests by bootstrapping pure AR processes, then there is no need for the econometrics toolbox. The Signal Processing Toolbox is needed for smoothing. If smoothing is done using a Bartlet window, then the toolbox is no longer necessary

9.3 List of functions

The following functions are available.

1. Folder **Auxiliary**

    - AROLS - AR model of a time series based on Ordinary Least Squares.
    - FourierSpectrum - Parametric estimate of the Fourier Power Spectrum of a time series, by fitting an ARMA process.
    - GMWMeasures - Some measures associated with a given generalized Morse wavelet.
    - MatrixMax - Local maxima of a matrix.
    - ProcessMatrix - Pre-processing of columns of given matrix.
    - SurrogateARMABoot - Surrogate series based on ARMA model and bootstrap.
    - SurrogateARMAEcon - Surrogate series using Econometrics toolbox.
    - WaveletSpectra - Wavelet transforms of all the columns of a given matrix.

2. Folder **WaveletTransforms**

    - AWT - Analytic wavelet transform of given series.
    - AWTOutput - Different quantities computed from a wavelet transform.
    - AWCO - wavelet coherency and cross-wavelet transform of two series.
    - AWCOOutput - Different quantities computed from a wavelet coherency.
    - MPAWCO - Multiple and partial wavelet coherencies.
    - MPAWCOOutput - Different quantities computed from multiple and partial wavelet coherencies.

References

[1] Aguiar-Conraria, L., Azevedo, N., Soares, M. J. (2008). Using wavelets to decompose the time-frequency effects of monetary policy. *Physica A: Statistical Mechanics and its Applications* 387:2863–2878.

[2] Aguiar-Conraria, L., Soares, M. J. (2011a). Oil and the macroeconomy: using wavelets to analyze old issues. *Empirical Economics* forthcoming, doi: 10.1007/s00181-010-0371-x.

[3] Aguiar-Conraria, L., Soares, M. J. (2011b). Business cycle synchronization and the Euro: a wavelet analysis. *Journal of Macroeconomics* forthcoming

[4] Baumeister, C, Peersman, G (2008). Time-varying effects of oil supply shocks on the US economy. *Gent University Working Paper 515*.

[5] Baubeau, P., Cazelles, B. (2009). French economic cycles: a wavelet analysis of French retrospective GNP series. *Cliometrica* 3:275–300.

[6] Baxter, M., King, R. (1999). Measuring business cycles: Approximate band-pass filters for economic time series, *The Review of Economics and Statistics* 81:575–593.

[7] Berkowitz, J., Kilian, L. (2000). Recent developments in bootstrapping time series. *Econometric Reviews* 19: 1–48.

[8] Blanchard, O. and Simon, J. (2001), The long and large decline in U.S. output volatility, *Brookings Papers on Economic Activity* 1: 135-64.

[9] Brémaud, P. (2002). Mathematical principles of signal processing: Fourier and wavelet Analysis. Springer-Verlag.

[10] Brooks, R., Del Negro, M. (2004). The rise in comovement across national stock markets: market integration or IT bubble? *Journal of Empirical Finance* 11: 659–680.

[11] Cazelles, B., Chavez, M., Magny, G., Guégan, J-F and Hales, S. (2007), Time-dependent spectral analysis of epidemiological time-series with wavelets. *Journal of the Royal Society Interface* 4: 625–36.

[12] Ciner, C. (2001), Energy shock and financial market: Nonlinear linkages. *Studies in Nonlinear Dynamics and Econometrics* 5: 203–212.

[13] Chatterjee, S. (1986). Bootstrapping ARMA models: Some simulations. *IEEE Transactions on Systems, Man and Cybernetics* 16: 294–299.

[14] Christiano, L. J., Fitzgerald, T. J. (2003). The band pass filter. *International Economic Review* 44: 435–65

[15] Connor, J., Rossiter, R. (2005). Wavelet transforms and commodity prices. *Studies in Nonlinear Dynamics and Econometrics* 9(1): Article 6.

[16] Craigmile, P., Whitcher, B. (2004). Multivariate spectral analysis using Hilbert wavelet pairs. *International Journal of Wavelets, Multiresolution and Information Processing* 2: 567–587.

[17] Crowley, P. (2007). A guide to wavelets for economists. *Journal of Economic Surveys* 21:207-267.

[18] Crowley, P. (2010). Long cycles in growth: explorations using new frequency domain techniques with US data. *Bank of Finland Research Discussion Papers 6/ 2010.*

[19] Crowley, P., Mayes, D. (2008). How fused is the Euro area core?: An evaluation of growth cycle co-movement and synchronization using wavelet analysis. *Journal of Business Cycle Measurement Analysis* 4: 63-95.

[20] Crowley P, Mayes D, Maraun D (2006a) How hard is the Euro area core? An evaluation of growth cycles using wavelet analysis. *Bank of Finland Research Discussion Paper No. 18/2006.*

[21] Crowley P, Mayes D, Maraun D (2006b) Analysing productivity cycles in the Euro area, US and UK using wavelet analysis. Available at http://www.business.curtin.edu.au/files/crowley_maraun_mayes.pdf.

[22] Daubechies, I. (1992). Ten Lectures on wavelets., volume 61 of CBMS-NSF Regional Conference Series in Applied Mathematics. SIAM, Philadelphia.

[23] Delprat, N., Escudié, B. , Guillemain, P. Kronland-Martinet, R., Tchamitchian, P., Torrésani, D. (1992). Asymptotic wavelet and Gabor analysis: Extraction of instantaneous frequencies. *IEEE Transactions on Information Theory* 38: 644–665.

[24] Farge, M. (1992). Wavelet transforms and their applications to turbulence. *Annual Review of Fluid Mechanics* 24: 395–457.

[25] Fernandez, V. (2005). The international CAPM and a wavelet-based decomposition of value at risk. *Studies in Nonlinear Dynamics and Econometrics* 9(4): Article 4.

[26] Forbes, K. and Rigobon, R. (2002). No contagion, only interdependence: Measuring stock market co-movements. *Journal of Finance* 57: 2223–2261.

[27] Foufoula-Georgiou, E., Kumar, P. (1994). Wavelets in Geophysics. In Wavelet Analysis and Its Applications, volume 4. Academic Press.

[28] Gallegati, M. (2008). Wavelet analysis of stock returns and aggregate economic activity. *Computational Statistics and Data Analysis 52*: 3061–3074.

[29] Gallegati, M. Gallegati, M. (2007). Wavelet variance analysis of output in G-7 countries. *Studies in Nonlinear Dynamics and Econometrics* 11(3): Article 6.

[30] Ge, Z. (2007). Significance tests for the wavelet power and the wavelet power spectrum. *Annals of Geophysics* 25: 2259–2269.

[31] Ge, Z. (2008). Significance tests for the wavelet cross spectrum and wavelet linear coherence. *Annals of Geophysics* 26: 3819–3829.

[32] Gençay, R., Selçuk, F., Withcher, B. (2001a). Scaling properties of foreign exchange volatility. *Physica A: Statistical Mechanics and its Applications* 289: 249–266.

[33] Gençay, R., Selçuk, F., Withcher, B. (2001b). Differentiating intraday seasonalities through wavelet multi-scaling. *Physica A: Statistical Mechanics and its Applications* 289: 543–556.

[34] Gençay, R., Selçuk, F., Withcher, B. (2005). Multiscale systematic risk. *Journal of International Money and Finance* 24: 55–70.

[35] Goupillaud, P., Grossman, A., Morlet, J. (1984). Cycle-octave and related transforms in seismic signal analysis. *Geoexploration* 23: 85–102.

[36] Grinsted, A., Moore, J. C., Jevrejeva, S.. Application of the cross wavelet transform and wavelet coherence to geophysical time series. *Nonlinear Processes in Geophysics* 11: 561–566.

[37] Holschneider, M. (1995). Wavelets: An analysis tool. Clarendon Press, Oxford.

[38] Huang, R., Masulis, R. and Stoll, H. (1996). Energy shocks and financial markets. *Journal of Futures Markets* 16: 1-17.

[39] Hudgins, L., Friehe, C., Mayer, M. (1993). Wavelet transforms and atmospheric turbulence. *Physics Review Letters* 71:3279–3282.

[40] Jagrič, T., Ovin, R. (2004). Method of analyzing business cycles in a transition economy: The case of Slovenia. *The Developing Economies* 42: 42–62.

[41] Kilian, L. (2009). Not all oil price shocks are alike: disentangling demand and supply shocks in the crude oil market. *American Economic Review* 99: 1053-1069.

[42] Kilian, L., Park, C. (2009). The impact of oil price shocks on the U.S. stock market. *International Economic Review* 50: 1267–1287.

[43] Kim, C.-J., Nelson, C. (1999). Has the U.S. Economy become more stable? A Bayesian approach based on a Markov-switching model of the business cycle. *Review of Economics and Statistics* 81: 608–616.

[44] King, M., Sentana, E., Wadhwani, S. (1992). Volatility and links between national stock markets. *Econometrica* 62: 901–933.

[45] King, M., Wadhwani, S. (1990). Transmission of volatility between stock markets. *Review of Financial Studies* 3: 5–33.

[46] Lachowicz, P. (2009). Wavelet analysis: a new significance test for signals dominated by intrinsic red-noise variability. *arXiv:0906.4176v1.*

[47] Lilly, J., Olhede, S. (2009). Higher-order properties of analytic wavelets. *IEEE Transactions on Signal Processing* 57: 146–160.

[48] Lilly, J., Olhede, S. (2010). On the analytic wavelet transform. *IEEE Transactions on Information Theory* 56: 4135–4156.

[49] Mallat, S. (1998). A Wavelet tour of signal processing. Academic Press, New York.

[50] Maraun, D., Kurths, J., Holschneider, M.(2007). Nonstationary Gaussian processes in wavelet domain: Synthesis, estimation, and significance testing. *Physical Review E* 75, 016707: 1–14.

[51] McConnell, M., Pérez-Quirós, G. (2000). Output fluctuations in the United States: What has changed since the early 1980s? *American Economic Review* 90: 1464–1476.

[52] Meyers, S. D., Kelly, B. G., O'Brien, J. (1993). An introduction to wavelet analysis in oceanography and meteorology: With application to the dispersion of Yanai Waves. *Monthly Weather Review* 121: 2858-2866.

[53] Olhede, S., Walden, A. (2002) Generalized Morse wavelets.*IEEE Transactions on Signal Processing* 50:2661–2670.

[54] Priestley, M. (1992). Spectral analysis and time series, 7th Edition, Academic Press, San Diego, CA.

[55] Ramsey, J. B. (1999). The contribution of wavelets to the analysis of economic and financial data. *Philosophical Transactions of the Royal Society of London Series A* 357: 2593–2606.

[56] Ramsey, J. B. (2002). Wavelets in economics and finance: past and future. *Studies in Nonlinear Dynamics and Econometrics* 6(3): Article1.

[57] Ramsey, J., Lampart, C. (1998a), Decomposition of economic relationships by time scale using wavelets: Money and income, *Macroeconomic Dynamics* 2: 49–71.

[58] Ramsey, J., Lampart, C. (1998b), The decomposition of economic relationships by time scale using wavelets: Expenditure and income, *Studies in Nonlinear Dynamics and Econometrics* 3: 23–42.

[59] Raihan, S , Wen, Y., Zeng, B. (2005). Wavelet: A new tool for business cycle analysis. *Working Paper 2005-050A, Federal Reserve Bank of St. Louis.*

[60] Rua, A.(2010). Measuring comovement in the time-frequency space. *Journal of Macroeconomics* 32: 685-691.

[61] Rua, A., Nunes, L. C. (2009). International comovement of stock market returns: A wavelet analysis. *Journal of Empirical Finance* 16: 632–639.

[62] Sadorsky, P. (1999), Oil price shocks and stock market activity. *Energy Economics*, 21: 449–469.

[63] Selesnick, I., Baraniuk, R., Kinsbury, N. (2005). The dual-tree complex wavelet transform. *IEEE Signal Processing Magazine* 22: 123–151.

[64] Torrence, C., Compo, G. (1998). A practical guide to wavelet analysis. *Bulletin of the American Meteorological Society* 79: 61–78.

[65] Torrence, C., Webster, P.(1999). Interdecadal changes in the ENSO-monsoon system. *Journal of Climate* 12: 2679–2690.

[66] Tu, C.-L., Hwang, W.-L., Ho, J. (2005). Analysis of singularities from modulus maxima of complex wavelets. *IEEE Transactions on Information Theory*, 51:1049–1062.

[67] Whitcher, B., Guttorp, P., and Percival, D. (2000). Wavelet analysis of covariance with application to atmospheric time series. *Journal of Geophysical Research* 105: 14941–14962.

[68] Wong, H., Ip, W.-C. , Xie, Z., Lui, X. (2003). Modelling and forecasting by wavelets, and the application to exchange rates. *Journal of Applied Statistics* 30: 537–553.

[69] Yogo, M. (2008). Measuring business cycles: A wavelet analysis of economic time series. *Economics Letters* 100: 208–212.