



# Toward new multi-wavelets: associated filters and algorithms. Part I: theoretical framework and investigation of biomedical signals, ECG, and coronavirus cases

Malika Jallouli<sup>1</sup> · Makerem Zemni<sup>1</sup> · Anouar Ben Mabrouk<sup>2,3,4</sup>  · Mohamed Ali Mahjoub<sup>1</sup>

Accepted: 27 August 2021 / Published online: 6 September 2021

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2021

## Abstract

Biosignals are nowadays important subjects for scientific researches from both theory, and applications, especially, with the appearance of new pandemics threatening the humanity such as the new coronavirus. One aim in the present work is to prove that wavelets may be a successful machinery to understand such phenomena by applying a step forward extension of wavelets to multi-wavelets. We proposed in a first step to improve multi-wavelet notion by constructing more general families using independent components for multi-scaling and multi-wavelet mother functions. A special multi-wavelet is then introduced, continuous, and discrete multi-wavelet transforms are associated, as well as new filters, and algorithms of decomposition, and reconstruction. Applied breakthroughs of the paper may be summarized in three aims. In a first direction, an approximation (reconstruction) of a classical (stationary, periodic) example dealing with Fourier modes has been conducted in order to confirm the efficiency of the HSch multi-wavelets in approximating such signals and in providing fast algorithms. The second experimentation is concerned with the decomposition and reconstruction application of the HSch multi-wavelet on an ECG signal. The last experimentation is concerned with a de-noising application on a strain of coronavirus signal permitting to localize approximately the transmembrane segments of such a series as neighborhoods of the local maxima of an numerized version of the strain. Accuracy of the method has been evaluated by means of error estimates and statistical tests.

**Keywords** Wavelets · Multi-wavelets · Wavelet filters · Wavelet algorithms · ECG · Coronavirus

✉ Anouar Ben Mabrouk  
anouar.benmabrouk@fsm.rnu.tn

Malika Jallouli  
jallouli.malika3@gmail.com

Makerem Zemni  
makaremzemni@gmail.com

Mohamed Ali Mahjoub  
medalimahjoub@gmail.com

<sup>1</sup> LATIS, Laboratory of Advanced Technology, and Intelligent Systems, Université de Sousse, Ecole Nationale d'Ingénieurs de Sousse, 4023 Sousse, Tunisia

<sup>2</sup> Laboratory of Algebra, Number Theory, and Nonlinear Analysis LR15ES18, Department of Mathematics, Faculty of Sciences, University of Monastir, 5000 Monastir, Tunisia

<sup>3</sup> Department of Mathematics, Higher Institute of Applied Mathematics, and Computer Science, Street of Assad Ibn Alfourat, University of Kairouan, 3100 Kairouan, Tunisia

<sup>4</sup> Department of Mathematics, College of Sciences, University of Tabuk, Tabuk, Saudi Arabia

## 1 Introduction and motivations

Signal/image processing studies digital signals/images, and their transformations, with the aim of improving their quality, extracting information, description, analyzing, and interpreting, control, filtering, compression, and transmission of data, de-noising, prediction, identification, and classification. These aims are met in many fields such as medicine, sound, geography, town planning, photography.

In theory, image/signal processing is a domain of science that is not recent, but in contrast, it is developed till the early discovery of Fourier analysis, and generally, linear transformations. See for example Bacchelli et al. (2002), Kotas and Moron (2017), Mallat (2008), and Wang et al. (2016).

It consists of a box of techniques, and/or methods, mathematical, and/or physical, theoretical, and/or applied that aims to modify or to convert a signal/image in another form in order to improve it, and/or to extract information (See Kotas and Moron 2017; Mallat 2008; Wang et al. 2016).

One of the important types of signals/images that attract researchers are biomedical ones. It is, for example, worthy to recall the pandemics, which appear in many periods somehow suddenly, and which cause a real threatening for humanity. Corona-type pandemics are one of them. They therefore need to be understood, such as SARS, H2N2, and the new coronavirus COVID-19.

One of the powerful tools in such topics is wavelet theory, which has been proved to be challenging since its discovery. Recently, a step forward has been also conducted to extend wavelets to multi-wavelets (another face of wavelets) to improve the theme of wavelet theory and its applications such as in signal processing. One of our aims in the present work is to improve multi-wavelet notion by adopting more general families of explicit multi-wavelets constructed using independent components for scaling functions. Readers may also consult (Arfaoui et al. 2020a,b; Ben Mabrouk et al. 2008b; Jallouli et al. 2019a,b,c,d; Zemni et al. 2019a) for more applications of the concepts of wavelets, and multi-wavelets, especially, on bio-signals.

Multimedia documents constitute also a category of applications in signal/image processing. They also present essential tools in the fields of biomedical, satellite, and astronomical imagery, film production, cryptography, watermarking, steganography, etc. In watermarking, for example, some methods operate in the spatial or transformed domain, and others use hybrid techniques. Concerning the frequency transform, the insertion of the mark in the low frequencies generally provides good robustness but induces distortions in the time domain. On the other hand, the insertion in the high-frequency components does not keep the quality, and moreover it makes the mark fragile to the attacks. This was a motivation for researchers to develop and ensure a compromise between the robustness, and invisibility of the transform. See Arfaoui et al. (2017), Bacchelli et al. (2002), Kotas and Moron (2017), Mallat (2008) and Wang et al. (2016).

Over the past few years, there has been a renewed interest in wavelet/multi-wavelet multi-resolution methods. When it is sought to analyze an image, it is very common to establish, explicitly or implicitly, a time-frequency representation of it. The Fourier transform is not the appropriate tool to carry out this analysis since it masks the temporal evolution of the signal. Wavelet theory has been proved to be a powerful tool in signal/image processing. Indeed, most of the signals of the real world are not stationary, and it is just in the evolution of their characteristics (statistics, frequency, temporal, spatial) that resides the essential information. In this context, wavelet transforms provide information about the frequency content, while preserving the localization in time in order to obtain a time-frequency representation or a space-scale of the signal. Unlike the Fourier transform, the wavelet transform provides interesting solutions in this context. Appoxima-

tions of signals are obtained as results of a convolution with a scaling function (a low-pass filter), and a wavelet function, and then reducing the number of points used in the process. The principle idea is to iterate this process and transform the current approximation into a new one with fewer points for the representation. We obtain a temporal as well as a frequency decomposition of the source object. It is well known that the frequency decomposition of a signal is interesting for the analysis of the different levels of detail present in the signal. It also applies to filtering, compression, and progressive transmission. See Arfaoui et al. (2017), Bacchelli et al. (2002), Daubechies (1992), Kotas and Moron (2017), Mallat (2008) and Wang et al. (2016).

Multi-wavelets are introduced to generalize wavelets to more flexible systems. In the original constructions, multi-wavelets start by exploiting the 2-scale relation due to the scaling function of a single multi-resolution analysis, by taking the well-known 2-scale relation in a vector form. Each component is a translated copy of the single source scaling function appearing in the 2-scale relation. More precisely, let  $\varphi$  be a scaling function satisfying an associated 2-scale relation, with filter length  $L$ . The associated multi-scaling function is  $\Phi(\cdot) = (\varphi(\cdot), \varphi(\cdot - 1), \dots, \varphi(\cdot - L + 1))$ . Original multi-wavelets look like a system of  $L$  surveillance systems in each direction, but which are identical or having the same mechanism in all directions. However, it will be best, and more efficient to install different mechanisms' cameras, and thus get a whole system of surveillance  $\Phi(\cdot) = (\varphi_1, \varphi_2, \dots, \varphi_K)$ , with a number of directional-wise cameras with different filters, independent, and working simultaneously to compose a whole image. In the present work, one of our aims is to apply the last mechanism of multiple different mechanisms.

Multi-wavelets are in fact another face of wavelets, looking like the multi-surveillance systems. They are vector-valued wavelets, satisfying more flexibility as single wavelets by involving matrix theory rather than scalars. This makes an important advantage, as it permits to obtain multi-wavelet bases possessing several properties at the same time. Matrix theory may help in obtaining an appropriate matrix product filter, spectral factorization, and in the improvement of the accuracy of the computed factors. The main properties are resumed essentially in orthogonality, symmetry, short support, high number of vanishing moments. The crucial point that may be raised in multi-wavelet coefficients may be the choice of a good prefilter which can provide a good approximation of the true initial coefficient sequences, when applied to the input data (See for instance (Cotronei et al. 1998; Cotronei and Puccio 1997; Geronimo et al. 1994; Ho 2002; Xia et al. 1996)). The number of high-amplitude wavelet coefficients created by a brutal transition like an edge is proportional to the width of the supports of the filters. For a more accurate localization of singularities, the number

of high-amplitude wavelet coefficients produced should be as small as possible. So, the supports of the filters should be as short as possible. Moreover, the more the vanishing moments, the smaller the coefficients can be produced over smooth regions at fine scales. Therefore, the multi-wavelet coefficients that belong to the noise component can be more easily distinguished at fine scales. The support size increases proportionally to the number of vanishing moments, and multi-wavelets can provide a better trade-off of this.

Geronimo, Hardin, and Massopust applied fractal interpolation to constructed multi-wavelets, that have shown next good characteristics, and have been applied widely. These multi-wavelets have a principal common point with ours, in the fact that they have independent components, and did not re-exploit the same translations of a single wavelet/scaling function to obtain a multi-case. See Geronimo et al. (1994), Ho (2002) and Xia et al. (1996).

In the present work, we propose to serve from explicit multi-wavelets already introduced in Zemni et al. (2019a), and next applied in Zemni et al. (2019b) to improve firstly the theoretical findings, and in modeling biomedical signals. The existing idea consists in a simple change in the well-known 2-scale relation by writing it in a vector form. This makes almost all existing constructions of multi-wavelets to look-like as modified representations of the same original wavelets. See AlMahamdya and Riley (2014), Alramahi et al. (2018), Alwan (2014), Attakitmongcol et al. (2001), Massopust et al. (1996), Rieder et al. (1996) and Turcajova (1999). In our work, based on the well-known wavelets of Haar, and Faber–Schauder we developed a simple variant of multi-wavelets that are not issued from one source, as in existing works. Haar, and Schauder explicit functions are applied in our case. This choice permits exact computations of necessary coefficients applied in the processing. They also permit to reduce the number of such coefficients and obtain the next generations recursively. However, we recall that other examples of multi-wavelets may be also obtained even explicitly by applying other scaling functions, and/or wavelet mothers different from the present case. Some interesting cases may be found in Arfaoui et al. (2020a, b). See also Bui and Chen (1998), Huang and Li (2011), Iyer (2001), Kessler (2009), Liang et al. (1996), Ruedin (2002), Selesnick (1998, 1999, 2000), Tham et al. (2000), Vehel and Aldroubi (1997), Xia (1998), Xia and Jiang (1999) and Xia et al. (1996) for more methods, and applications.

Next, to show the performance of our extension, some experimentation will be developed. A first one deals with the development of a Fourier type mode to show how fast are algorithms based on the new variant of multi-wavelets. A second experimentation will be concerned with the well-known ECG signals. A de-noising step has to be conducted using our new multi-wavelets to lead next to a good analysis.

See AlMahamdya and Riley (2014), Kotas and Moron (2017) and Wang et al. (2016) for some existing methods.

The last experimentation is concerned with the processing of a coronavirus strain for an associated membrane protein signal. We propose to develop a wavelet analysis of an isolated or purified strain of human coronavirus associated with SARS already recorded, and studied in Van Der Werf (2010). Precisely, we intend to conduct a decomposition process, and to localize the transmembrane helices (TMHs) of the strain based on the hydrophobic character of the amino acids constituting the proteins' series associated to such a strain, and issued from the well-known Kyte–Doolittle method (Kyte and Doolittle 1982). The idea lies in the topic of molecular or cellular communication and its modeling by means of signals. Recall that the functioning of our body, as well as its interaction with exterior factors such as viruses is in fact a kind of molecular communication. For example, viruses respond to signalling molecules secreted to discover the exterior space. This is the simplest way to describe the mechanism of an attack. In neuronal system, electrical impulses and neurotransmitters are jointly used by neuron cells to communicate with target cells. The question resides on how the communication is conducted. Recall that viruses are infectious agents, that may not replicate solely, and in contrast, they need a hosting living cell to assure the replication inside it. Some of them have an own protein cover, and all of them develop proteins inside the host cell. This is necessary to assure and to permit the communication with other cells. We know that the communication is assured by means of membrane proteins. More precisely, transmembrane proteins are the main factors or parts in the body that permit the communication between cells. Our idea here has two meanings, aims or interpretations. In the first, the transmembrane proteins of the virus may be used by the virus itself to attack other cells, and in the same time, they may be considered in an inverse problem as weak parts, and open doors in the protein cover of the virus to receive itself exterior attacks such as those due to the body immune system, and those due to vaccines. So, in our opinion, all the problem resides in localizing these strong attackers, and/or these weak defenders. The first work applying wavelets as part of proteins modeling is due to Fischer et al. (2003), where a WAVPRED prediction algorithm of transmembrane segments as maximum points in a numerical series converted from Kyte–Doolittle scales, and which consists of mathematical and biochemical filtering by taking empirically determined filtering thresholds.

The present paper is organized as follows. The next section is devoted to the review of wavelet theory. Section 3 is devoted to the development of multi-wavelets in order to provide a Haar–Schauder multi-wavelet, and its associated filters. Recall that the original simple way to introduce multi-wavelets is to consider multi-wavelet scaling function as the vector composed of the translated copies of the same single

scaling function appearing in the 2-scale relation. Here, a different concept of multi-wavelet scaling functions is introduced based on finitely many possibly independent scaling functions components. The multi-scaling function and its multi-wavelet mother will be vectors whose components are not issued from the same single scaling function or scaling wavelet. It looks like a system of many cameras working simultaneously, and independently to provide a complete surveillance system. In Sect. 4, some experiments have been developed to show the performance of multi-wavelets against wavelets for both the rapidity of algorithms and bio-signals processing. An ECG signal, and a proteins' strain issued from a coronavirus case are considered.

## 2 Signal processing techniques review

Signal processing techniques are resumed in the various transformations, especially those based on mathematical concepts, and which are implemented using numerical, and/or digital techniques. They permit to characterize systems' processes in a quantified way, to reveal hidden information about the process, and account for the system behavior, and also it allows us to predict this behavior when the system's condition changes.

The complexity of bio-physical processes such as ECG, DNA, and proteins needs a quantified way that relies on the use of mathematical, and physical models and laws to understand them. However, there are no universal techniques, which may be appropriate for all applications and also, it is not practically possible to express all signal processing methods for various applications.

There are in the literature many techniques, such as the entropy, filtering, frequency analysis, time series analysis and models, such as auto regressive, and autoregressive moving average, spectral methods, like power spectral density, wavelets, thresholding, denoising, Fourier transform, Hilbert–Huang transform, uncertainty principle, support vector machine, adaptive noise cancellation used to enhance the signal-to-noise ratio.

Signal denoising has been conducted also via the empirical mode decomposition, which is based on a local and adaptive method in the frequency-time analysis. Some others are based on statistical model such as the so-called deep-learning-based autoencoder models. These models consist in regenerating a clean version of the analyzed signal from a corrupted version based on an optimization process of a suitable objective function. In the same category, Bayesian filters based models such as the extended Kalman filter, the extended Kalman smoother, and the unscented Kalman Filter are also known in bio-signals processing. Fuzzy models are also applied widely, and also combined with neural networks for signal processing. Besides, we may also mention

the hybrid methods developed by combining these ones. See Abhijith et al. (2016), AlMahamdy and Riley (2014) and Babatunde (2012)

Support vector machines, for example, are used in signal processing to separate different patterns by means of pattern recognition. The technique is based on statistical learning theory made by learning from the collected set of data. This method is also combined with wavelets to yield the wavelet support vector machine, which has been widely applied. See for instance (Abhijith et al. 2016; AlMahamdy and Riley 2014; Alwan 2014; Babatunde 2012; Ben Mabrouk et al. 2008a; Ho 2002; Mallat 2008; Xia and Suter 1996; Zemni et al. 2019a, b).

Heisenberg's uncertainty principle, for example, states that it is not possible to know what specific frequency exists at a particular instance of time but it is only possible to know what frequency bands exist at what time interval. The problem of time and frequency resolution which is the result of the Heisenberg uncertainty principle constitutes a major challenge in the analysis of non-stationary signals. The use of wavelets in such a principle is nowadays a very well-known fact. Time series models, linear and nonlinear, are also applied in signal processing as approximate mathematical model based on sets of inputs–outputs measurements.

Wavelet transform, which is the main technique related to the present work, is applied to signals to obtain further information that is not readily obtainable in the raw signal. Most signals in practice are time-domain signal in their raw format. Moreover, almost all biological signals are non-stationary signals. Wavelet transform is capable of simultaneously providing both the time and frequency information. The frequency spectrum represents the frequency components of a signal. Fourier transform is used to find the frequency-amplitude representation of a signal. However, many signals such as ECG and proteins need more than the Fourier transform theorems, due to their frequency contents which may change in time. Wavelet analysis is capable of revealing aspects of data that other signal analysis techniques cannot, e.g., trends, breakdown points, discontinuities in higher derivatives and self-similarity.

Wavelet methods start firstly by decomposing the signal, deciding the type of thresholding and reconstructing the signal. Recently, wavelets have been extended to multi-wavelets, which have shown some performance compared to existing methods. Methods based on wavelets/multi-wavelets include the mathematical theory of irregular functions to conduct signal processing, such as the estimation of Lipschitz exponent by means of wavelet/multi-wavelet coefficients or transform, which performs the singularity detection, and thus yields signal denoising algorithms using the singularity detection. A thresholding process permits to de-noise the signal, and reconstruct the denoised version by simply applying the inverse multi-wavelet transform. In this direc-

tion, Geronimo, Hardin and Massopust proposed a method for constructing translation and dilation invariant function spaces using fractal functions defined by a certain class of iterated function systems (Geronimo et al. 1994). Xia and collaborators (Xia 1998; Xia and Jiang 1999; Xia et al. 1996; Xia and Suter 1996) improved such multi-wavelets by constructing a prefilter design method dealing with all decomposition steps for the discrete multi-wavelet transform, and approximating a signal with the lowpass property. See also Attakitmongcol et al. (2001), Bacchelli et al. (2002), Bui and Chen (1998), Cotronei et al. (1998); Cotronei and Puccio (1997), Cotronei and Sissouno (2017), Efromovich (2001), Hardin and Roach (1998), Ho (2002), Ho et al. (2003), Johnson (2000), Kotas and Moron (2017), Strela et al. (1999), Wang et al. (2016) and Yoganand and Mohan (2018).

### 3 Wavelets/multi-wavelets for signal processing

#### 3.1 Wavelet methods review

A wavelet may be defined simply as a short wave function, and which has major difference from Fourier sine, and cosine by its ability of being localized in time-frequency, and/or time-space. Wavelet analysis of signals is based on the so-called wavelet transform which is a convolution of the analyzed signal with copies of a source function called mother wavelet. Wavelets, differently from Fourier modes, are not necessarily periodic, they may be also compactly supported.

In mathematics, a mother wavelet  $\psi$  is a square-integrable function with enough vanishing moments (oscillating) with necessary zero mean. Such a mother wavelet has to satisfy some admissibility assumption, stating that

$$\mathcal{A}_\psi = \int_{\mathbb{R}} \frac{|\widehat{\psi}(\xi)|^2}{|\xi|} d\xi < \infty. \tag{1}$$

(See Arfaoui et al. 2017; Daubechies 1992; Mallat 2008). The copies applied next in the signal analysis are issued from the mother wavelet by translation, and dilation parameters. More precisely, the wavelet processing of signals is based on their wavelet transform. Given a finite energy signal  $F$ ,  $a > 0$  known as the scale, and  $b \in \mathbb{R}$  known as the position, the continuous wavelet transform (CWT) of  $F$  is at the scale  $a$ , and the position  $b$  is

$$C_F(a, b) = \int_{-\infty}^{+\infty} F(t)\psi_{a,b}(t)dt, \tag{2}$$

where

$$\psi_{a,b}(x) = \frac{1}{\sqrt{a}}\psi\left(\frac{x-b}{a}\right). \tag{3}$$

The analyzed signal  $F$  may be reconstructed using the inverse transform as

$$F(t) = \frac{1}{\mathcal{A}_\psi} \int_{-\infty}^{+\infty} C_F(a, b)\psi_{a,b}(t)\frac{dadb}{a^2}, \tag{4}$$

where  $\mathcal{A}_\psi$  is the admissibility constant due to the mother wavelet  $\psi$  defined by (1). (See Arfaoui et al. 2017; Daubechies 1992; Mallat 2008).

A restrictive version of the CWT is the so-called discrete wavelet transform (DWT) called also wavelet coefficient, evaluated by the restriction to discrete grids for the scale, and the position parameters. In fact there is no essential difference between the discrete grids used. The most commonly used one is the dyadic grid constituted by  $a = 2^{-j}$ , and  $b = k2^{-j}$ ,  $j, k \in \mathbb{Z}$ . The copy  $\psi_{a,b}$  becomes is this case

$$\psi_{j,k}(t) = 2^{j/2}\psi(2^j t - k) \tag{5}$$

and the discrete wavelet transform (DWT), called sometimes, the wavelet coefficient, will be

$$d_{j,k}(F) = \int_{-\infty}^{+\infty} F(t)\psi_{j,k}(t)dt. \tag{6}$$

These coefficients are also known in wavelet theory as the detail coefficients at the level  $j$ , and the position  $k$ . It holds also in wavelet theory that  $(\psi_{j,k})_{j,k \in \mathbb{Z}}$  constitutes an orthonormal basis of  $L^2(\mathbb{R})$ , and consequently any element  $F$  may be decomposed in a series

$$F = \sum_{j,k} d_{j,k}(F)\psi_{j,k} \tag{7}$$

known as the wavelet series of  $F$ , and which replaces the reconstruction formula (4) in the discrete form.

This decomposition into an orthogonal-wise components series leads to a functional framework associated to the mother wavelet  $\psi$  known as the multi-resolution analysis (MRA). Indeed, let for  $j \in \mathbb{Z}$ ,  $W_j = \text{spann}(\psi_{j,k}; k \in \mathbb{Z})$  known as the detail spaces, and  $V_j = \bigoplus_{l \leq j} W_l$  called approximation spaces. There exists a source function  $\varphi$  known as the scaling function or the father wavelet satisfying  $V_j = \text{spann}(\varphi_{j,k}; k \in \mathbb{Z})$ , where the  $\varphi_{j,k}$ 's are defined similarly to the  $\psi_{j,k}$ . The father, and mother wavelets are related by the so-called 2-scale relation stating that

$$\varphi = \sum_{k \in \mathbb{Z}} h_k \varphi_{1,k}, \quad \text{and} \quad \psi = \sum_{k \in \mathbb{Z}} g_k \varphi_{1,k}, \tag{8}$$

where

$$h_k = \int_{-\infty}^{+\infty} \varphi(t)\varphi_{1,k}(t)dt, \quad \text{and} \quad g_k = (-1)^k h_{1-k}. \quad (9)$$

See Daubechies (1992) and Mallat (2008) for more details. These relations permit to compute the wavelet coefficients from level to level. Indeed, denote

$$a_{j,k}(F) = \int_{-\infty}^{+\infty} F(t)\varphi_{j,k}(t)dt,$$

known as the approximation or the scaling coefficient of  $F$  at the level  $j$ , and the position  $k$ , we have

$$a_{j,k}(F) = \sum_{l \in \mathbb{Z}} h_l a_{j+1,l+2k}(F), \quad (10)$$

and

$$d_{j,k}(F) = \sum_{l \in \mathbb{Z}} g_l a_{j+1,l+2k}(F). \quad (11)$$

This means that the decomposition at the level  $j$  may be deduced from the level  $(j + 1)$  by means of the filters  $H = (h_k)_k$  (discrete wavelet low-pass filter), and  $G = (g_k)_k$  (discrete wavelet high-pass filter). Similarly, we have an inverse scheme stating that

$$a_{j+1,k}(F) = \sum_l h_{l-2k} a_{j,l}(F) + \sum_l g_{l-2k} d_{j,l}(F). \quad (12)$$

For backgrounds on wavelet filters, the readers may refer to Arfaoui et al. (2017), Daubechies (1992) and Mallat (2008).

In wavelet theory, the series (7) may be decomposed into two parts,

$$F = \sum_{j \leq J_0, k} d_{j,k}(F)\psi_{j,k} + \sum_{j > J_0, k} d_{j,k}(F)\psi_{j,k}, \quad (13)$$

where  $J \in \mathbb{Z}$ . Due to the properties of the multi-resolution analysis, the first part above belongs in fact to the so-called approximation space of level  $J$  denote usually  $V_J = \text{spann}(\varphi_{J,k}, k)$ , and consequently may be expressed by means of the  $\varphi_{J,k}$ 's as

$$A_{J_0}(F) = \sum_{j \leq J_0, k} d_{j,k}(F)\psi_{j,k} = \sum_k a_{J_0,k}(F)\varphi_{j,k}, \quad (14)$$

where the  $a_{J,k}(F)$  are the approximation coefficients of  $F$  introduced above.  $A_{J_0}(F)$  is effectively called the approximation of  $F$  at the level  $J_0$ , which is also the projection of

$F$  on  $V_{J_0}$ . The second part is a superposition of orthogonal components

$$D_j(F) = \sum_k d_{j,k}(F)\psi_{j,k} \quad (15)$$

in the so-called detail space  $W_j = \text{spann}(\psi_{j,k}, k)$  of the multiresolution analysis.  $D_j(F)$  is effectively called the detail component of  $F$  at the level  $j$ , which is also the projection of  $F$  on  $W_j$ . In other words, we may write

$$F = A_{J_0}(F) + D_{J_0+1}(F) + D_{J_0+2}(F) + \dots \quad (16)$$

It is composed of a first part describing the global behavior or the shape of  $F$ , and a second part reflecting the higher frequency oscillations or the fine scale deviations of the series near its trend. In practice, we cannot obviously compute infinitely many parts, but we fix a maximal level of decomposition  $J > J_0$ , and consider

$$F_J = A_{J_0}(F) + \sum_{J_0 < j \leq J} D_j(F). \quad (17)$$

There is no theoretical method for the exact choice of the parameters  $J_0$  and  $J$ . However, the minimal parameter  $J_0$  does not have an important effect on the total decomposition and usually chosen to be 0. But, the choice of  $J$  is always critical. One selects  $J$  related to the error estimates.

### 3.2 Multi-wavelet processing

Multi-wavelets have been introduced since the early 1990s as another view of wavelets permitting to re-write wavelet analysis in a vector form to reduce may be mathematical formulations. It resembles in some sense to the reduction of higher-order differential equations into first order ones by considering the vector  $X = (y, y', y'', \dots, y^{(n)})$  where  $n \in \mathbb{N}$  is an integer constituting the order of the original differential equation in  $y$ , and where  $y', y'', \dots, y^{(n)}$  are the derivatives of  $y$  to such an order.

The major existing multi-wavelet constructions consider the vector  $\Phi = (\varphi(\cdot), \varphi(\cdot - 1), \dots, \varphi(\cdot - N))$ , where  $N$  is the length of the filters  $H$ , and  $G$ . This view of wavelets has even though some advantages, such as short supports, smoothness, accuracy, symmetry, and orthogonality. Moreover, as noticed in Zemni et al. (2019a, b), discrete multi-wavelets may require pre-processing, and post-processing steps. These facts themselves constituted main motivations behind the study developed in Zemni et al. (2019a, b), and continued in the present paper.

Some original developments of multi-wavelets have been already addressed by many authors, mainly (Geronimo et al. 1994). Geromino et al abbreviated in the literature as GHM

scaling functions have four remarkable properties showing that multi-wavelets can combine more useful features than scalar wavelets: both scaling functions have short supports. The system has second order of approximation. The translates of the scaling functions and wavelets are orthogonal. Both scaling functions and the wavelets are symmetric. The GHM multiscaling and multi-wavelet functions are also quite smooth, and precisely, almost differentiable.

Xia (1998) proposed a prefilter design by combining the ideas of the conventional wavelet transforms and multi-wavelet transforms. The prefilters are orthogonal but non-maximally decimated. The author stated that one benefit of such construction is that the energy compaction ratio with the GHM multi-wavelets compared to Daubechies wavelets. See also Cotronei et al. (1998); Cotronei and Puccio (1997), Xia (1998), Xia and Jiang (1999) and Xia et al. (1996).

More about mutiwavelets and their applications may be found in Efromovich (2001), Fowler and Hua (2002), Johnson (2000), Lebrun and Vetterli (1998), Lebrun and Vetterli (2001), Ruedin (2003), Shen and Tan (2001), Stacey and Blyth (2008), Strela et al. (1999), Yoganand and Mohan (2018).

In the present paper, we propose to continue in exploiting more the construction of multi-wavelets as noticed in Zemni et al. (2019a, b) by considering a vector-valued scaling function  $\Phi = (\varphi_1, \varphi_2, \dots, \varphi_N)^T$  ( $T$  is the transpose),  $N \in \mathbb{N}$  fixed, where the components  $\varphi_i, i = 1, 2, \dots, N$  are not translations of the same function as in the most existing cases. This leads to a matrix-vector 2-scale relation

$$\Phi = \sum_k H_k \Phi_{1,k}, \tag{18}$$

where in this way the  $H_k$ 's are  $(N, N)$ -matrices,  $H_k = (h_{i,j})_{1 \leq i,j \leq N}$ . Similarly, the mother multi-wavelet will satisfy a scale relation of the form

$$\Psi = \sum_k G_k \Phi_{1,k}, \tag{19}$$

where the coefficients  $G_k$ 's are also  $(N, N)$ -matrices,  $G_k = (g_{i,j})_{1 \leq i,j \leq N}$ .

**Definition 1** (Zemni et al. 2019a) The sequences of matrices  $H = (H_k)_k$ , and  $G = (G_k)_k$  are called the discrete high pass, and discrete low pass multi-filters, respectively.

In the literature review on multi-wavelets, there are few developments. So, complete, and full exposition of multi-wavelets theory still needs to be developed. Only few references in this direction are known such as Cotronei et al. (1998); Cotronei and Puccio (1997), Keinert (2004), Xia (1998), Xia and Jiang (1999) and Xia et al. (1996). This is one motivation among previous ones letting us to develop the present work. The

choice of mother multi-wavelets is also strongly related to the ability, and flexibility in conducting experiments. Readers may refer to Attakitmongcol et al. (2001), Brazile (2009), Hardin and Roach (1998), Keinert (2004), Stankovic and Falkowski (2003), Zhang et al. (2001) and Wang et al. (2016)

In the sequel, we fix the multi-wavelet order  $N = 2$ . Let  $\varphi_1(x) = \chi_{[0,1]}(x)$  be the Haar scaling function, and  $\varphi_2(x) = (1 - |x|)\chi_{[-1,1]}(x)$  be the Schauder scaling function. Denote next  $\Phi = (\varphi_1, \varphi_2)^T$ . Simple calculus yield that  $H_k = 0$  whenever  $|k| \geq 2$ , and

$$\Phi = H_{-1}\Phi_{1,-1} + H_0\Phi_{1,0} + H_1\Phi_{1,1} \tag{20}$$

where

$$\begin{aligned} H_{-1} = H_1 &= \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 0 \\ 0 & 1/2 \end{pmatrix}, \\ H_0 &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \end{aligned} \tag{21}$$

Thus, the mother multi-wavelet is

$$\Psi = \sum_l G_l \Phi_{1,l}, \quad G_l = (-1)^l H_{1-l}. \tag{22}$$

The Haar–Schauder multi-wavelet processing (decomposition/reconstruction) of a signal  $F$  consists as in all wavelet processing in estimating the corresponding coefficients of the signal by means of the multi-wavelet copies. So, consider, for  $r \in \mathbb{N}$  fixed, known as the order or the dimension of the signal, a signal  $F = (F_1, F_2, \dots, F_r)^T$ . Denote also  $A_{j,k}(F)$ , and  $D_{j,k}(F)$  the approximation, and the detail coefficients of  $F$ , relatively, to the Haar–Schauder multi-wavelets at the level  $j$ , and the position  $k$ . The signal  $F$  may be decomposed as a sum

$$F = A_0 + D_0$$

where  $F_0$  is

$$A_0 = \sum_l A_{0,l}(F)\Phi_{0,l} \tag{23}$$

and

$$D_0 = \sum_l D_{0,l}(F)\Psi_{0,l}. \tag{24}$$

The components  $A_0$ , and  $D_0$  are known as the approximation, and the detail components of  $F$  at the level 0. The coefficients  $A_{0,l}(F)$ , and  $D_{0,l}(F)$  are  $(r, 2)$ -matrices. As in the case of single wavelet theory, we obtain here a MRA associated to the multi-wavelet by considering as approximation space  $V_0$  the closure of vector space spanned by the  $\Phi_{0,l}$ , and as detail

space at the level 0 the one spanned by  $\Psi_{0,l}$ ,  $l \in \mathbb{Z}$ . As a consequence, we obtain multi-wavelet algorithms stating that

$$A_{1,s}(F) = \sum_l [A_{0,l}(F)H_{s-2l} + D_{0,l}(F)G_{s-2l}], \quad (25)$$

$$A_{0,s}(F) = \sum_l H_{l+2s}A_{1,l}(F) \quad (26)$$

and

$$D_{0,s}(F) = \sum_l G_{l+2s}A_{1,l}(F). \quad (27)$$

In general, as in the case of single wavelet theory, we get here a multi-wavelet decomposition of the signal  $F$  at the level  $J$  ( $J_0$  being chosen to be equal to 0.) as

$$F^J = AS_J + \sum_{j=0}^J DS_j. \quad (28)$$

To resume, the new general concept will cover some disadvantages of many existing multi-wavelets theory where the scaling multi-wavelet function is constructed by taking the well-known 2-scale relation in single wavelet theory, and introducing the multi-wavelet scaling function as the vector composed of the translated copies of the single source scaling function appearing in the 2-scale relation.

### 3.3 Wavelets/multi-wavelets brief comparison

Compared to single wavelets, multi-wavelets have many advantages due to these characteristics. Owing more than one scaling functions, multi-wavelets permit to use correct stencils, and to identify the low and high frequency efficiently. The property of vanishing moments maintains convergence of higher order upto boundaries. Moreover, multi-wavelets permit the reduction of computational overhead. The use of a set of short support filters in multi-wavelet leads to dual benefits over scalar wavelets. The first one is that multi-wavelet with a given support can achieve the smoothness offered by scalar wavelets with larger support. The second benefit is that multi-wavelet provides better compaction than the scalar wavelets. Furthermore, multi-wavelets have the advantage that the user can optimize the multi-wavelet system for any application.

In fact, many techniques have been developed for biomedical signals. We may mention the data compression, such as amplitude-zone-time epoch coding, the coordinate reduction time coding system, turning point technique, prediction, modulation, and also transformational methods such as Fourier, Walsh, Karhunen-Loeve, Wavelet and multi-wavelets.

Wavelet transform has been shown to be an efficient tool in signal processing aimed at compressing ECG signals, detection of QRS complex, analysis of ventricular late potential, localizing knots in DNA and proteins' series, prediction of anomalies, etc. See Ben Mabrouk and Ibrahim Mahmoud (2013), Fischer et al. (2003) and Ibrahim Mahmoud et al. (2016)

Wavelets/multi-wavelets are also applied for noise elimination by adopting singularity detection for example. Ho et al in Ho et al. (2003) proposed a multi-wavelet method for singularity detection for regularity scalable image coding. See also Ho (2002).

Finally, based on the literature and ideas above, finding a suitable multi-wavelet(s) for bio-signals processing is of interest. Haar, and Faber–Schauder, when combined yield a multi-wavelet that possesses the majority of properties for multi-wavelets. They also permit to well approximate simultaneously piece-wise, and linear cases.

## 4 Experimentation

In this section, the wavelet/multi-wavelet method will be applied for processing three examples. In the first part, a simple example consisting of a Fourier mode estimation will be provided. Next, an ECG signal will be considered. Finally, a special example dealing with a coronavirus modeling will be developed.

The idea consists in using the HSch multi-wavelet for the ECG signal processing as a type of simultaneous loops to guarantee the maximum information carried in such a signal. The first loop consists in applying a filtering of the signal by means of one of the components of the HSch multi-wavelet (Haar for example), and next apply the second one to denoise more the obtained filtered sub-signal. This raises an interesting question about the use of independent components in the definition of the multi-wavelet analysis source functions  $\Phi$ , and  $\Psi$ . This filtering concept could not be realized by using multi-wavelets with non-separable variables, and/or dependent components. So, the idea is a double (multiple in general) surveillance cameras system that is used to detect best the strange objects.

We now explain mathematically the principle of HSch multi-wavelet processing. So, denote  $\varphi$ , and  $\tilde{\varphi}$  the Haar, and Faber–Schauder scaling functions, respectively, and the associated mother wavelets  $\psi$ , and  $\tilde{\psi}$ . For a level  $J$  denote  $a_J$ , and  $\tilde{a}_J$  the approximations at the level  $J$  due to Haar, and Faber–Schauder MRA, respectively, and similarly  $d_J$ , and  $\tilde{d}_J$  the projections on the detail spaces due to Haar, and Faber–Schauder MRA, respectively. We get the multi-wavelet decomposition of the ECG signal at the level  $J$  as

$$A_J = a_J + \tilde{a}_J + \sum_j^J d_j + \sum_j^J \tilde{d}_j.$$

Using the independence between the components of the multi-wavelet, the principle applied here means that the final decomposition is a superposition of two decompositions on two approximation spaces, and two detail spaces for each level included in the modeling. In this case, the risk of losing the information decreases compared with classical wavelet processing. The reconstruction by multi-wavelet will be more efficient. Moreover, it is worth to recall here that there is no essential difference between being simultaneous or consecutive the application of the two components of the multi-wavelet. Such a problem may be of great importance when the components are dependent or depending on non-separable variables.

The following diagram in Fig. 2 illustrates more the decomposition steps of signals using the Hsch multi-wavelet. Besides, Fig. 1 illustrates the principle of the multi-wavelet decomposition, and reconstruction of signals, as well as the computation of the error of approximation of the signal by wavelets. Algorithm 1 shows the headlines of the computer code for such decomposition/reconstruction.

**Algorithm 1** ECG Reconstruction

```

Require: OriginalECG
Ensure: ReconstECG
    NQ1 ← 1
    NQ2 ← 0
    J ← 1
    N ← 1
    while N ≠ 0 do
        [HDetail, HApp, SchDetail, SchApp] ←
            HSCH – decompositionJ(OriginalECG)
        ReconstECG ← HSCHReconstructionJ
            (HDetail, HApp, SchDetail, SchApp)
        NQ2 ← NAQE(ECGReconst)
        if NQ2 > NQ1 then
            N ← 0
        else
            J ← J + 1
            NQ1 ← NQ2
        end if
    end while
    
```

Finally, to illustrate the closeness of the wavelet/multi-wavelet method approximation, we propose to compute the Normalized Average Quadratic Error (NAQE) to show the performance of the approximation computed on a grid of  $N$  points  $t_i$  in the time domain of the time series  $X$ , as

$$NAQE(X, X_a) = \frac{\sum_{i=1}^N (X(t_i) - X_a(t_i))^2}{\sum_{i=1}^N (X(t_i))^2}, \tag{29}$$

where  $X_a$  is the corresponding approximation of  $X$  relative to the method used.

**4.1 Development of a Fourier mode**

In this section, we aim to develop the multi-wavelet analysis of a simple example of signals consisting of the well known  $2\pi$ -periodic Fourier mode  $F(t) = \sin(t), t \in [0, 2\pi]$ . The purpose is essentially to provide a simple example that may be re-conducted by readers, and to show based on this example the performance, and the superiority of multi-wavelets in the processing, especially, the multi-wavelets composed of different components that are not translations of the same single scaling function or mother wavelet. This example is essentially characterized by being periodic, with no singularity. These characteristics makes its approximation (reconstruction) to be well conducted by multi-wavelets as in the case of the coincidence between the Fourier series and the analyzed function when all the well assumption of the Dirichlet theorem are satisfied. Recall that the decomposition de  $F$  at the level  $J \in \mathbb{N}$  is expressed as

$$F = \sum_k A_{J,k}(F)\varphi_{J,k} + \sum_{j \geq J} \sum_k D_{j,k}(F)\psi_{j,k}. \tag{30}$$

For a choice of  $J = 1$ , the approximation part becomes

$$A_1 = \sum_k A_{1,k}(F)\varphi_{1,k}. \tag{31}$$

Recall now that

$$A_{1,k}(F) = \int_{(k-1)/2}^{(k+1)/2} \sin(t)\varphi_{1,k}(t)\chi_{[0,2\pi]}(t)dt. \tag{32}$$

We now compute the values of the position parameter  $k$  for which the intersection of supports  $[\frac{k-1}{2}, \frac{k+1}{2}] \cap [0, 2\pi] \neq \emptyset$  which yields that  $0 \leq k \leq [4\pi]$ .

We next compute the Normalized Average Quadratic Error (NAQE) to show the performance of the approximation computed on a grid of  $N$  points  $t_i$  in  $[0, 2\pi]$ ,

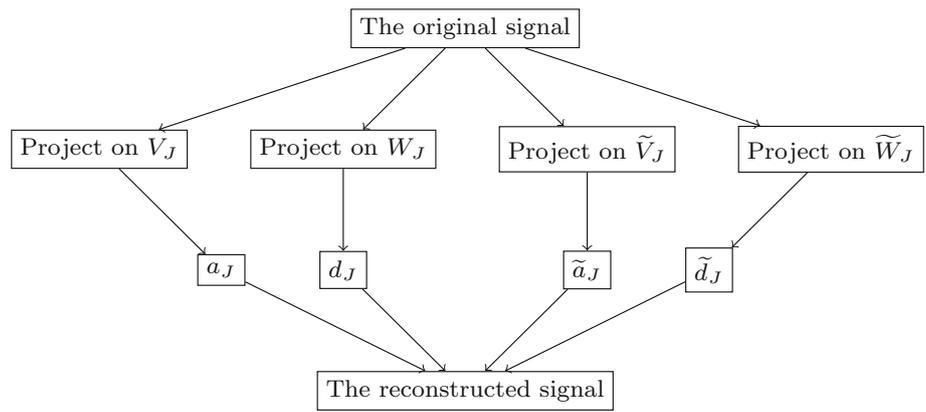
$$NAQE_{J,N}(A_1, F) = \frac{\sum_{i=1}^N (A_J(t_i) - F(t_i))^2}{\sum_{i=1}^N (F(t_i))^2}. \tag{33}$$

For a number  $N = 50$ , and  $J = 1$ , we get an error

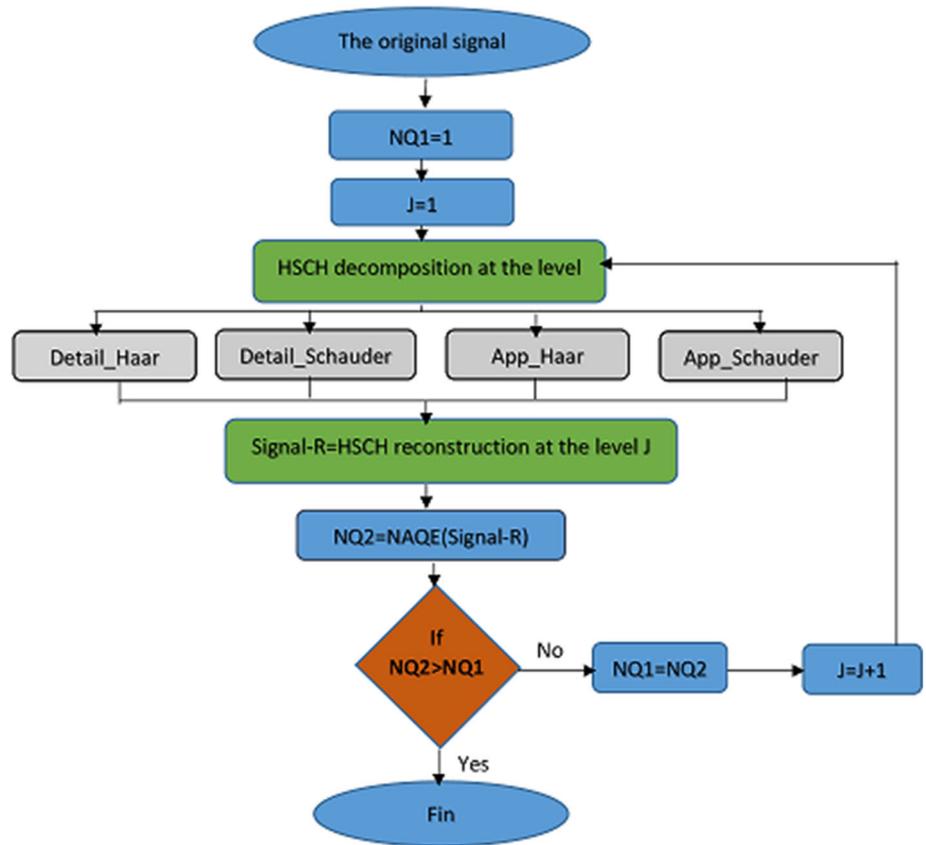
$$NAQE = 0.0012.$$

The following figure (Fig. 3) illustrates the signal  $F$ , and its approximation  $A_1$ .

**Fig. 1** The HSch multi-wavelet principle



**Fig. 2** Schematic illustration of the HSch multi-wavelet principle



Next, to show the role of the projections of the signal  $F$  on the detail spaces we compute the DWT of  $F$  already with  $J = 1$ . This will illustrate the dynamics of  $F$ . Similarly to the approximation case, it remains for the position parameter  $k$  the values  $-1, 0, 1, \dots, 13$ . Recall that the support of  $\psi_{1,k}$  is  $\left[ \frac{k-1/2}{2}, \frac{k+1/2}{2} \right]$ . Thus, to get the detail component  $D_1$  of the signal  $F$  in the detail space  $W_1$  we have to compute  $D_{1,k}(F)$  for  $k \in \{-1, 0, 1, \dots, 13\}$ .

Next, denote  $F_1 = A_1 + D_1$ . To illustrate the closeness of  $F_1$  to the original signal  $F$ , we compute as previously the

NAQE on a grid of  $N$  points  $t_i$  in  $[0, 2\pi]$ ,

$$NAQE_N(F_1, F) = \frac{\sum_{i=1}^N (F_1(t_i) - F(t_i))^2}{\sum_{i=1}^N (F(t_i))^2}. \tag{34}$$

For a number  $N = 50$ , we get an error

$$NAQE = 0.00118.$$

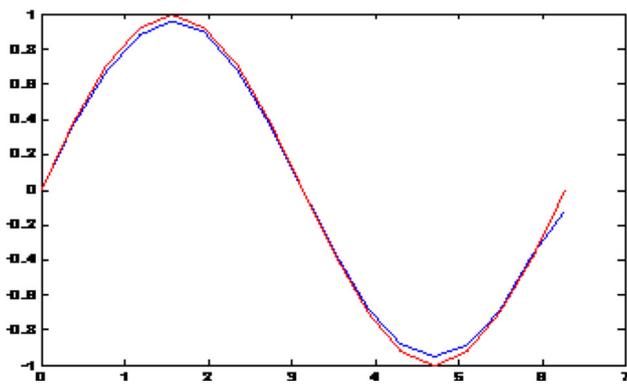


Fig. 3  $F$  (red), and its approximation  $A_1$  (blue) (color figure online)

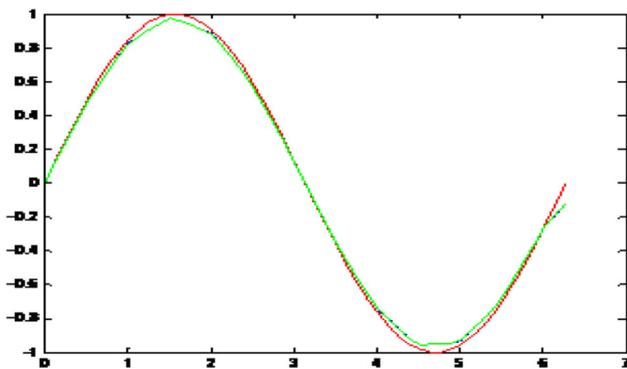


Fig. 4  $F$  (red), and  $F_1$  (green) (color figure online)

Table 1 Error estimates

The method	Corresponding error
Method 1 (Ref. Brazile 2009)	$12 \cdot 10^{-3}$
Bi-filters $J = 1$	$11,8 \cdot 10^{-3}$
Bi-filters $J = 2$	$2 \cdot 10^{-5}$
Bi-filters $J = 3$	$2,4 \cdot 10^{-6}$

The following figure (Fig. 4) illustrates the signal  $F$ , and its approximation  $F_1$ .

Similarly, we may compute for  $J \in \mathbb{N}$  the approximation

$$F_J = A_J + D_1 + D_2 + \dots + D_J. \tag{35}$$

To illustrate the closeness of these approximations to the original signal  $F$ , we compute the Normalized Average Quadratic Error (NAQE) on a grid of  $N$  points  $t_i$  in  $[0, 2\pi]$ . For a number  $N = 50$ , and  $J = 1$ , we get the following error estimates (Table 1).

Table 1 summarizes the results of comparisons with the existing method developed in Brazile (2009), and the bi-filters based method developed here. We found that NAQE obtained by bi-filters is smaller than the existing one. On the other hand, it is remarkable that the greater  $J$  increases the error decreases.

Table 2 Time execution

The method	NAQE	Running time
Schauder wavelet	0.0086	123.2 s
Schauder filters	0.0092	73.03 s
HSch multiwavelet	0.0033	32.97 s
HSch multi-wavelet filters	0.0003	16.6 s

Next, in order to show more the performance of the new method we proposed to evaluate the running time of algorithms due to each method. We thus provided a comparison relatively to the time execution algorithms for the methods applied for the same Fourier mode signal. For  $N = 10$ , and  $J = 1$ , we obtained the following table (Table 2).

Table 2 shows a comparison for both the NAQE error, and the execution time between the approximation obtained by the use of the Schauder wavelet, Schauder filters, HSch multi-wavelet, and HSch multi-wavelet filters. First, by comparing the NAQE, and the execution time for the methods based on the single Schauder wavelet, and Schauder filters we noticed that the NAQE relative to both of them are not enough different. Besides, the second one yields a faster convergent algorithm. Next, applying HSch multi-wavelets results in more efficient approach. Similarly to the single case, the new HSch multi-wavelet filters result in a best error, and a best running time. This shows the performance of the new multi-wavelet approach. Finally, our work proves among the efficiency of multi-wavelet approaches, that using different wavelet cells in the multi-wavelet black boxes is more performant than applying the classical approach. Recall that this latter is based on re-writing the 2-scale relation, and thus re-writing the whole signal in a different way by decomposing it in different multi-signals, which may affect the originality of the signal processed.

### 4.2 ECG signal processing

ECG signals are graphical representations of the heart electrical activity due to the variations of electric potential of the specialized cells in the contraction (myocytes), and specialized cells in the automatism, and the conduction of the influxes. ECG can highlight various cardiac abnormalities, and has an important place in cardiology diagnostic tests, as for coronary heart disease. We refer to the MIT-BIH Arrhythmia data basis for the application developed in this part.

Similarly to the last example, an estimation of the original ECG signal with its  $J$ -imation  $F_J$  defined by Eq. (35) is provided. The HSch multi-wavelet of the ECG signal may be written as

$$ECG_J = A_J(ECG) + D_1(ECG) + \dots + D_J(ECG), \tag{36}$$

**Table 3** Relative NAQE estimates for ECG signal

NAQE	H-W	Sch-W	HSch MW
$J = 1$	0.0012	0.0014	$9.8 \cdot 10^{-4}$
$J = 2$	$8.7 \cdot 10^{-4}$	$8.95 \cdot 10^{-4}$	$7.4 \cdot 10^{-4}$
$J = 3$	$5.46 \cdot 10^{-4}$	$6.18 \cdot 10^{-4}$	$4.37 \cdot 10^{-4}$
$J = 4$	$5.01 \cdot 10^{-4}$	$3.7 \cdot 10^{-4}$	$1.09 \cdot 10^{-4}$

where  $A_J(EGC)$  is the approximation at the level  $J$  of the ECG signal due to the HSch multi-wavelet obtained by projection the original signal ECG on the  $J$ -level approximation space  $V_J$  due to the HSch multi-wavelet multi-resolution. The  $D_j(EGC)$ ,  $1 \leq j \leq J$  are, respectively, the detail components of the ECG signal at the corresponding levels  $j$ , obtained as usual by projecting the original signal ECG on the corresponding detail spaces  $W_j$ ,  $1 \leq j \leq J$ , due to the HSch multi-wavelet multi-resolution. The closeness of the approximation  $ECG_J$  to the original signal ECG is evaluated via the NAQE error

$$NAQE_N(EGC, J) = \frac{\sum_{i=1}^N (ECG_J(t) - ECG(i))^2}{\sum_{i=1}^N (ECG(i))^2} \quad (37)$$

estimated on the time interval of the ECG signal. The present ECG signal processing by means of the HSch multi-wavelet yields for each level of decomposition  $J \geq 1$  a discrete positions' grid,  $0 \leq k \leq 10.2^J$ .

Table 3 resumes the accuracy of the present method against previous ones by means of the so-called Normalized Average Quadratic Error (NAQE) in (37).

This type of signals is one of the most complex cases in signal processing due to their high volatility (fluctuation), point-wise irregular from the mathematical point of view. These bad characteristics make their modeling and/or approximation to be delicate. In the present work, one aim is to show the performance of the multi-wavelet machine in overcoming this ambiguity by providing best estimation of these signals in few time of execution. We notice easily from Table 3 that the HSch multi-wavelet processing results in more accurate error of closeness NAQE obtained for the best estimates at a level of decomposition  $J = 4$ . This proves also that the multi-wavelet processing did not necessitate a higher order of decomposition to reach a good error. Besides, Figs. 5, 6, and 7 illustrate the processing of the ECG signal using Haar wavelet (H-W), Schauder wavelet (Sch-W), and HSch multi-wavelets (HSch-MW), and confirm more the efficiency, and the performance of the multi-wavelet principle.

To finish with the ECG multi-wavelet processing, we plotted in Fig. 8 the evolution of the normalized quadratic error with the level of resolution  $J$ , for Haar wavelet, Schauder wavelet, and HSch multi-wavelets. The graph shows easily the efficiency of wavelet processing in general, and more efficiently the dominance of the new multi-wavelet against the single wavelets.

### 4.3 A case of coronavirus signal

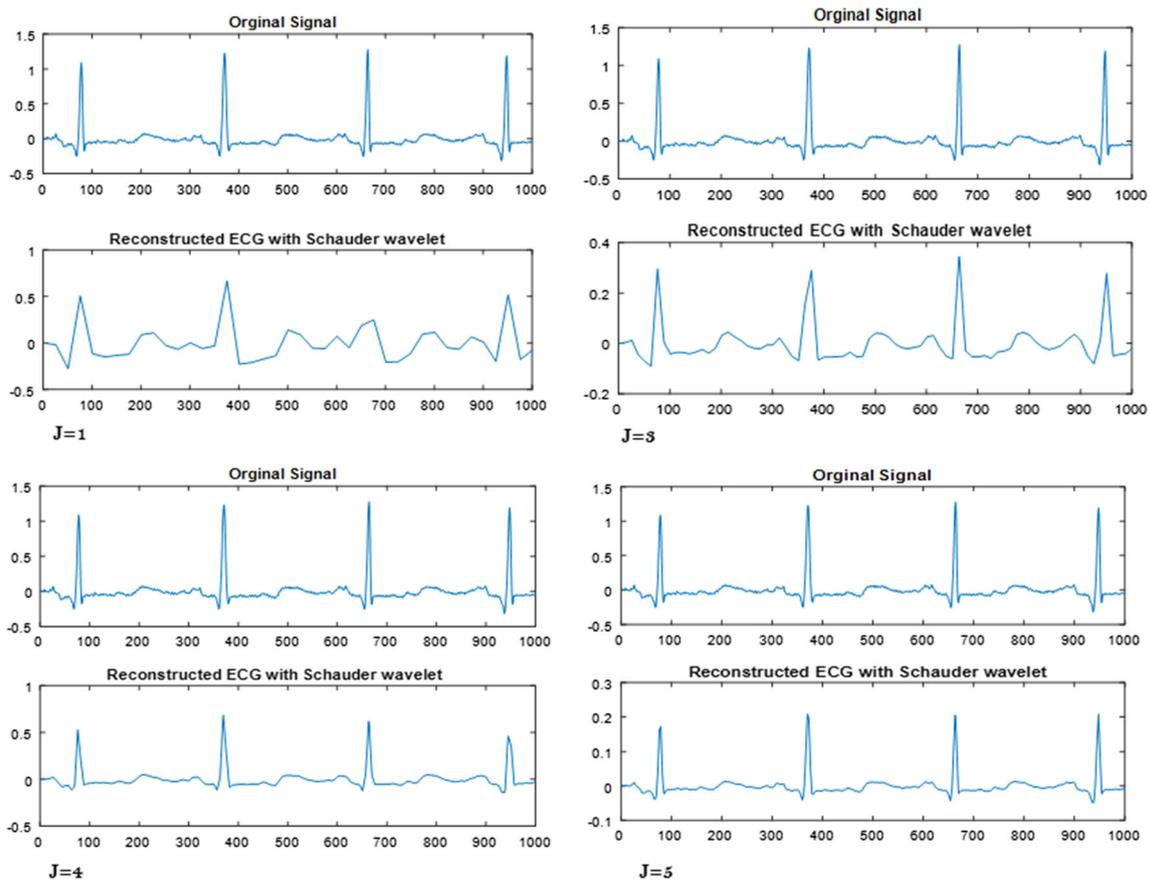
We consider in this work a strain of coronavirus associated with SARS, from a sample originally recorded in Hanoi, Vietnam since 2002–2003, See Van Der Werf (2010). Recall that the coronavirus is not indeed new, except that, it appears each time in a new form or a new state. It is for example enveloped, and includes, on its surface, peplomic structures called spicules. It may, and precisely always includes proteins of unknown encoded function. Such proteins have several categories. Some are, for example, membrane glycoproteins in the form of spicules emerging from the surface of the viral envelope. They are responsible for attaching the virus to receptors in the host cell, and for inducing fusion of the viral envelope with the cell membrane. Other proteins of even small variable sizes are transmembrane proteins. They play a crucial role in the budding process of coronaviruses which occurs at the level of the intermediate compartment in the endoplasmic reticulum, and the Golgi apparatus.

This makes the localization of transmembrane segments of great importance, and this is in fact the main purpose of this part of the paper; how to localize or detect these segments for the case of Coronavirus?

Indeed, the localization of such segments is important for the comprehension of the virus functioning, and mechanism, as transmembrane segments are the main attackers, which permit, and/or which are responsible of the exchange between the virus and the exterior space, such as the human body cells, and receptors.

Moreover, the localization, the comprehension of the placement of these segments on the whole strain, and their composition with proteins, will permit to practitioners such as doctors, and drug makers to best prepare their attacks against the virus.

Membrane proteins constitute more than the quarter of proteins in currently sequenced complete genomes. They have a very important role in cellular processes such as the transportation of molecules, and the communication between cells. Moreover, they are directly, and strongly related to drugs. More than the half of such proteins are targeted by a drug each one. Inside the membrane, the transmembrane segments may take the form of an alpha helix or the beta strand form. Generally, the size of the TM segments is of the order of 15–30 amino acids with a very large hydrophobic region.



**Fig. 5** Reconstruction of the ECG signal by Schauder wavelet

When infecting a host cell, the reading frame of the viral genome is translated into a polyprotein which is cleaved by viral proteases, and then releases several non-structural proteins such as RNA polymerase, and ATPase helicase. These two proteins are involved in the replication of the viral genome as well as in the generation of transcripts which are used in the synthesis of viral proteins.

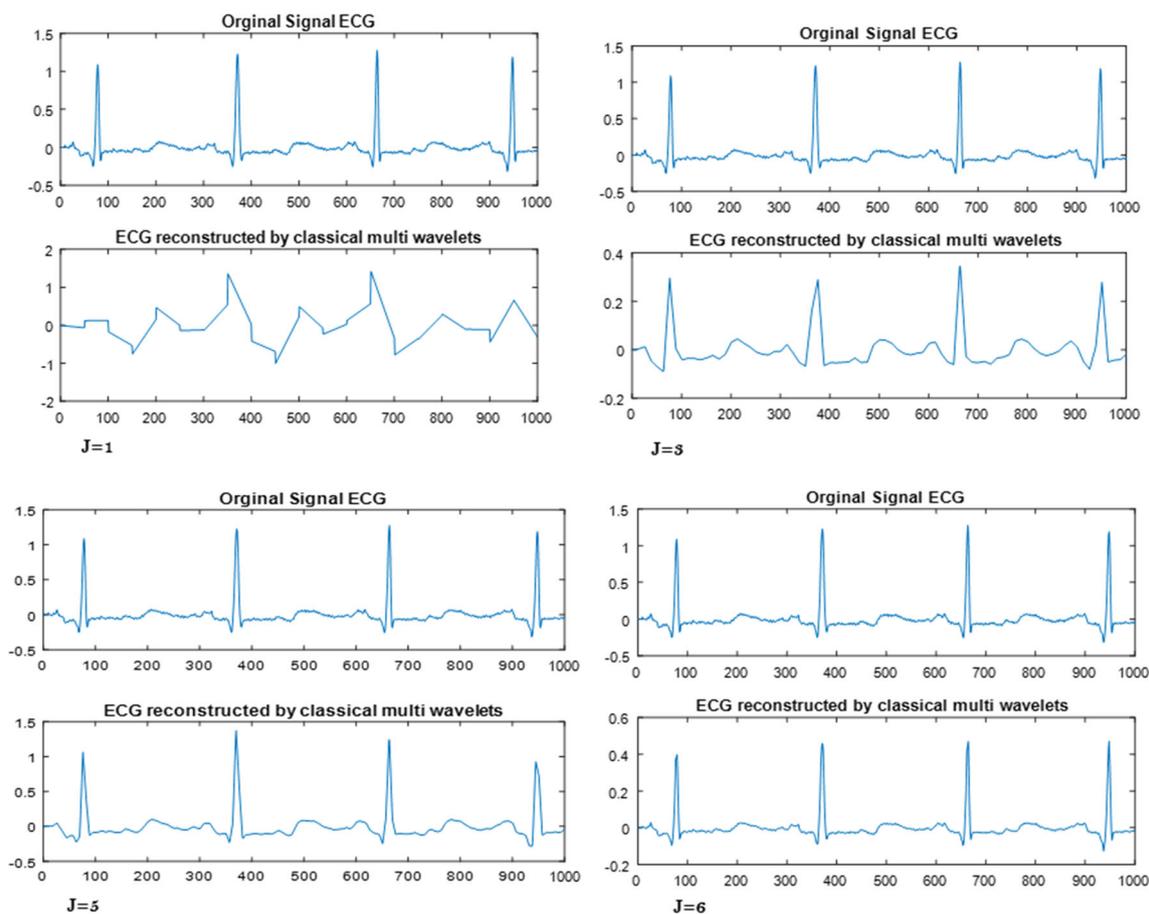
With the help of proteins, the virus migrates through the Golgi complex, and leaves the cell, and thus attaches to external bodies causing hard damages. Indeed, coronaviruses are responsible for humans, and animals, for colds, respiratory, and digestive infections by inducing antibodies.

The coronavirus appeared in several forms such as SARS which spread to different countries in 2002–2003. Very recently a new type of the same category of epidemics appeared originally in Hanoi, China, and presents until now a challenge for humanity. The severity of these diseases is the rate or the growth of mortality in the first place, and the auto-internal change of the virus although its external form appears the similar. Determining the causative agent of the new category is now the challenge for all of humanity. More information, and ideas on such type of viruses may be found in Anand et al. (2002), Bonnin (2018), Desjardins (2010), Li

(2016), Liu et al. (2006), McBride et al. (2014), Talbot and Jouvence (1992) and Xu et al. (2016).

The purpose of this work is to apply a wavelet/multi-wavelet analysis of an isolated or purified strain of human coronavirus associated with SARS already recorded, and studied in Van Der Werf (2010).

Recall that proteins' sequences are biological series similar, and also related to DNA as they are characters' series, and which also may be generated from DNA ones. The question of why preferring proteins, and not DNA as others do is already discussed in Zemni et al. (2019a). One main cause is due to the fact that proteins' sequences are more volatile. On the other hand, sequences of DNA are always issued from proteins' ones as for the example applied here. Moreover, the communication between living cells such as virus ones are always done by the intermediary of membrane, and precisely transmembrane proteins. The regions of anomalies, and communication constitute some type of helices which correspond to the singular, and optimum points in the numerical series issued from the biological ones. See Arfaoui et al. (2020a), Fischer et al. (2003), Ibrahim Mahmoud et al. (2016) and Zemni et al. (2019a) for more details.



**Fig. 6** Reconstruction of the ECG signal by Haar wavelet

Now, a natural question is how to state out the role of the prefiltering and the HSCh multi-wavelet and their relationship with the improved performance, especially when compared with the existing methods.

It is well known that many functions may be well approximated by means of special types of approximators such as the ones expressed by the projections on the multi-resolution spaces associated to precise wavelets.

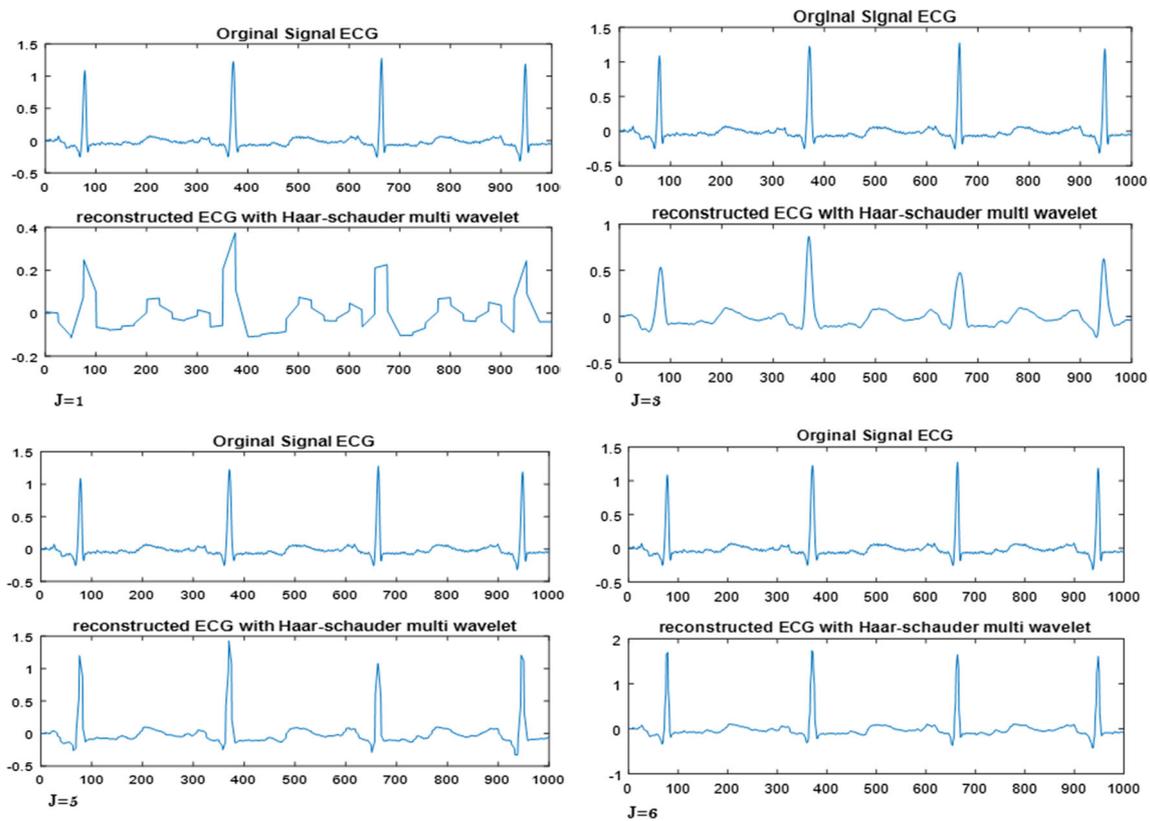
The oldest systems may be the Fourier one and next the Haar system. Although the Fourier decomposition leads to good results in many cases, some disadvantages are inherent such as the loss of the information concerning the space behavior of the signal. A discontinuity or a localized high variation of the frequency will not be well described by the Fourier representation. The underlying reason lies in the nature of complex exponential functions used as bases functions. They all cover the entire real line and differ only with respect to frequency. They are not suitable for representing the behavior of a discontinuous function or a signal with high localized oscillations.

For a Haar basis the approximation at a level  $J$  is piecewise constant. However, piecewise constant approximators

of smooth functions are far from optimal. A simple step ahead may be done by using piecewise linear approximators such as Faber–Schauder ones for approximating such functions.

One of the most close signals to piecewise constant and piecewise linear functions may be those obtained from the numerization methods due to biological series such as DNA and proteins. The obtained numerical time series are combinations of these two types of functions. One of the most commonly used methods in this field is the interpretation of hydrophathy profiles of protein series investigated in our work. This method was first introduced by Kyte and Doolittle (Babatunde 2012) who used a window of 19 residues to smooth the hydrophathy data, to enable the detection of potential transmembrane helices as peaks in a two-dimensional plot. Recall that a well-known problem in protein modeling is the prediction of the position of transmembrane helices (HTMs) in protein sequences. The window size is set to be 19 residues, due to the fact that most transmembrane elements are  $\alpha$ -helices about 18 residues long (Scarlati, <http://www.biophysics.org/btol/Scarlati.html>).

In the present work, one aim is to continue to explore wavelet methods to the prediction of transmembrane helices

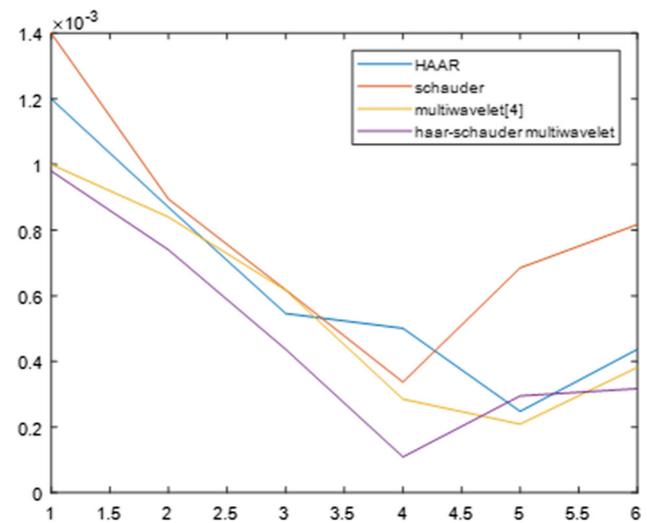


**Fig. 7** Reconstruction of the ECG signal by the HSCh multi-wavelet

using hydrophobicity scales. The main purpose is to improve the wavelet methods by multi-wavelet ones, and by applying precisely the closest multi-wavelets to the analysed series. This is the main reason of applying the HSCh multi-wavelet.

Indeed, the study and characterization of membrane proteins experimentally is a long, expensive method which requires advanced and well-specified equipment, which is why it is first necessary to extract the transmembrane proteins from their membrane environment. Also, it is essential to maintain them in their native functional and soluble form by reconstituting a host medium in aqueous solution similar a to that of the biological membrane. The extraction of these proteins is done using surfactants called membrane detergents, compounds capable of isolating, solubilizing and manipulating membrane proteins (Dauvergne 2010) as well as their crystallization, NMR spectroscopy, X-ray diffraction, etc. This is why we are trying to find a reliable method for the prediction of transmembrane helices based on mathematical and computer processing.

In this experimental part, a multi-wavelet process is developed to localize the transmembrane helices of the strain of the SARS-associated coronavirus based on the hydrophobic character of the amino acids developed in Kyte and Doolittle (1982). This permitted to convert proteins into time (numerical) series allowing their processing using mathematical tools



**Fig. 8** Error estimates relatively to the decomposition level  $J$  for ECG signal

to be possible (See Ben Mabrouk and Ibrahim Mahmoud 2013; Ben Mabrouk et al. 2015; Ibrahim Mahmoud et al. 2016). The numerical conversions due to Kyte–Doolittle (1982) are resumed in Table 4.

**Table 4** Hydrophobicity scale of Kyte-Doolittle

Amino acid	The scale	Category
Isoleucine: Ile(I)	+4.5	Hydrophobic
Valine: Val(V)	+4.2	Hydrophobic
Leucine: Leu(L)	+3.8	Hydrophobic
Phenylalanine: Phe(F)	+2.8	Hydrophobic
Cysteine: CySH(C)	+2.5	Hydrophobic
Methionine: Met(M)	+1.9	Hydrophobic
Alanine: Ala(A)	+1.8	Hydrophobic
Glycine: Gly(G)	-0.4	Neutral
Threonine: Thr(T)	-0.7	Neutral
Serine: Ser(S)	-0.8	Neutral
Tryptophan: Try(W)	-0.9	Neutral
Tyrosine: Tyr(Y)	-1.3	Neutral
Proline: Pro(P)	-1.6	Neutral
Histidine: His(H)	-3.2	Hydrophilic
Glutamine: Gln(Q)	-3.5	Hydrophilic
Asparagine: Asn(N)	-3.5	Hydrophilic
Glutamic Acid: Glu(E)	-3.5	Hydrophilic
Aspartic Acid: Asp(D)	-3.5	Hydrophilic
Lysine: Lys(K)	-3.9	Hydrophilic
Arginine: Arg(R)	-4.0	Hydrophilic

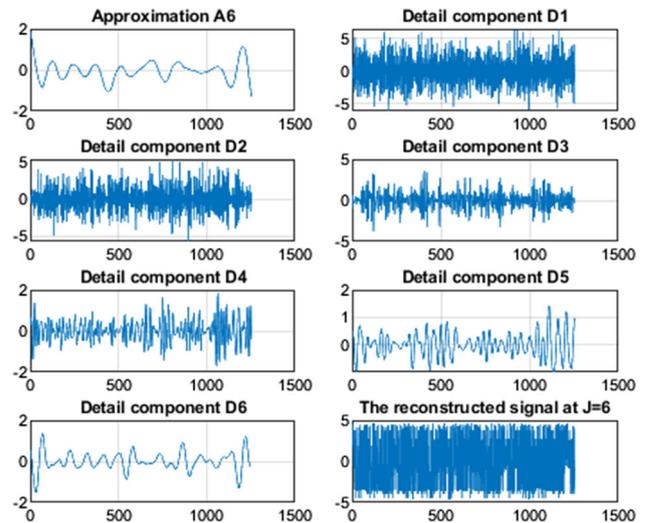
The protein strain is provided in Appendix 6. The multi-wavelet filtering acts as for the single wavelet case (or also the Fourier analysis) on the multi-wavelet (multi-scaling) coefficients of the analyzed signal by exploiting the 2-scale relation due to the HSch multi-scaling function (18) or (19). Let for  $J$  fixed  $A_J$  be the vector composed of all multi-scaling approximation coefficients  $A_{J,k}$ . For example, for  $J = 0$ , the approximation  $A_0$  will be observed as the vector  $A_0 = (A_{0,0}, A_{0,1}, \dots, A_{0,K_0^A})$ , obtained by the matrix form  $A_0 = M_0^\Phi X$ , where  $M_0^\Phi$  is the matrix whom coefficients are the blocks  $\Phi(i - j)$ ,  $0 \leq i \leq K_0^A$ ,  $1 \leq j \leq N$ , where  $N$  is the size of the original series  $X$ . Similarly, the vector  $D_0 = (D_{0,0}, D_{0,1}, \dots, D_{0,K_0^D})$  is evaluated by  $D_0 = M_0^\Psi X$ , where  $M_0^\Psi$  is the matrix whom coefficients are the blocks  $\Psi(i - j)$ ,  $0 \leq i \leq K_0^D$ ,  $1 \leq j \leq N$ . Theoretically speaking the dimensions  $K_0^A$ , and  $K_0^D$  of the approximation vector  $A_0$  and the detail  $D_0$  are different (or generally, the dimensions  $K_J^A$ , and  $K_J^D$  of the approximation vector  $A_J$  and the detail  $D_J$ ). However, to avoid the problem of dimension, we complete these vectors with zero coefficients to use one dimension. Now, for a level  $J$  fixed, the vector  $A_J$  and  $D_J$  will be divided into odd and even parts, for which we have

$$A_{J,2n} = A_{J-1,n}H_0 + A_{J-1,n-1}G_2 + A_{J-1,n}G_0,$$

$$A_{J,2n+1} = A_{J-1,n}H_1 + A_{J-1,n+1}H_{-1} + D_{J-1,n}G_1,$$

**Table 5** NAQE estimates for the coronavirus signal using HSch multi-wavelet

$J$	$NAQE_J$
1	$3.5108 \cdot 10^{-12}$
2	$4.3631 \cdot 10^{-12}$
3	$4.9646 \cdot 10^{-12}$
4	$5.2140 \cdot 10^{-12}$
5	$5.380 \cdot 10^{-12}$
6	$4.3631 \cdot 10^{-12}$



**Fig. 9** The decomposition of the numerized coronavirus proteins' series with HSch multi-wavelet at the level  $J = 6$

and similarly, the matrix form of the reconstruction will be expressed by means of the relations

$$A_{J-1,n} = H_{-1}A_{J,2n-1} + H_0A_{J,2n} + H_1A_{J,2n+1},$$

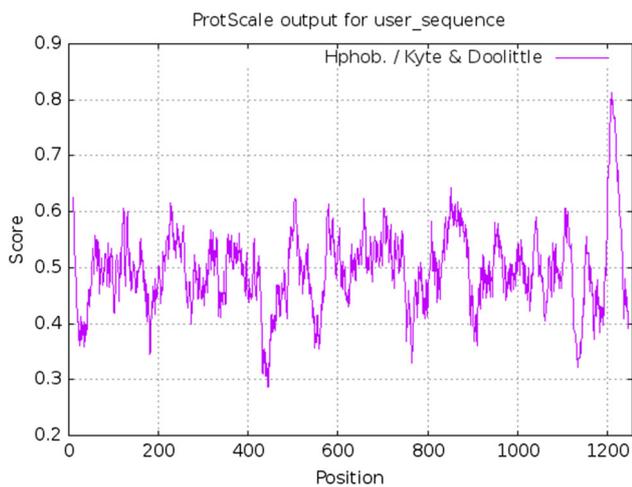
and

$$D_{J-1,n} = G_0D_{J,2n} + G_1D_{J,2n+1} + G_2D_{J,2n+2}.$$

To illustrate the closeness of the reconstructed signal to the original one, we computed as usual the NAQE. We get the estimates provided in Table 5.

Table 5 shows an optimal reconstruction reached at the level  $J = 6$ . Such optimality is explained by the fact that such a level is the minimum one from which the number of eventual transmembrane segments is stabilized at the number 8 segments. Next, Fig. 9 illustrates graphically the decomposition of the numerized coronavirus proteins' series at the level  $J = 6$  using HSch multi-wavelet. This shows in some part the efficiency of using multi-wavelets instead of single wavelets.

Next, as it is now well known that wavelets, and multi-wavelets are powerful tools to detect the transmembrane segments in proteins' series (Arfaoui et al. 2020a; Ben



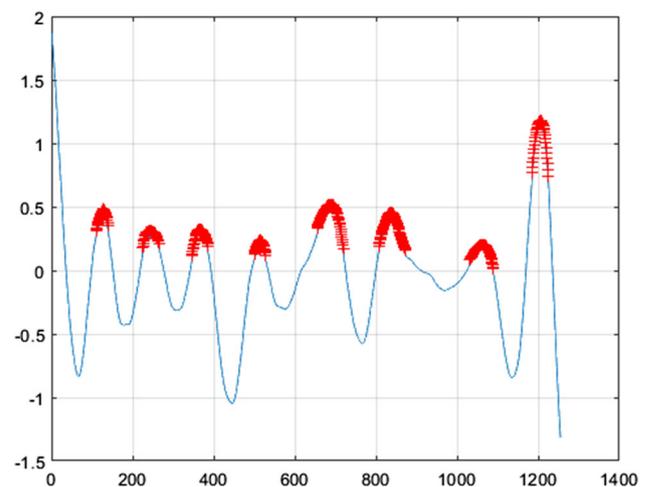
**Fig. 10** Kyte–Doolittle hydropathy signal for the coronavirus series

**Table 6** The TMHs Segments for HSCh filtering of the coronavirus signal

TMHs	HSCh multi-wavelet localized segments
1	120–134
2	233–253
3	359–373
4	505–523
5	678–699
6	824–842
7	1056–1069
8	1199–1212

Mabrouk and Ibrahim Mahmoud 2013; Ben Mabrouk et al. 2015; Ibrahim Mahmoud et al. 2016; Zemni et al. 2019a), and in order to prove the applicability, and thus the useful aspect of our multi-wavelet we proposed to focus on the possible detection, and/or prediction of alpha-helices in the considered protein. We subsequently propose to predict the locations of these regions by statistical processing applying the HSCh multi-wavelet. The optima with scores greater than 1.8 (horizontal line in Fig. 10) indicate possible transmembrane regions. The window position values shown on the  $x$ -axis of the graph reflect the average hydropathy of the entire window, with the corresponding amino acid as the middle element of that window. Eight helices (local maxima) appear clearly.

To show the efficiency of the present method, we apply next the new explicit HSCh multi-wavelet filtering at the optimal level  $J = 6$ . Table 6 illustrates the findings, and shows 8 segments. Next, we illustrated graphically such prediction in Fig. 11 which illustrates the predicted results due to the 'new' HSCh multi-wavelet at the level  $J = 6$ . It shows also 8 localized transmembrane helices.



**Fig. 11** TMHs prediction using HSCh multi-wavelet for the coronavirus signal

There are several statistical methods that may be applied to check the performance, accuracy and efficiency of the proposed methods and models investigated (Millett 2005). Generally, in the existing studies, it is stated that the predicted transmembrane helices are considered admissible if at least a half of them coincide with the observed (real) ones. The accuracy of the model/method is usually evaluated via the percentage index  $Q_p$  defined by Ben Mabrouk and Ibrahim Mahmoud (2013); Ben Mabrouk et al. (2015), Bin and Zhang (2013), Ibrahim Mahmoud et al. (2016) and Zemni et al. (2019a)

$$Q_p = \frac{N_{cor}}{\sqrt{N_{obs}N_{prd}}} \times 100\%,$$

where  $N_{cor}$  is the number of correctly predicted TMHs,  $N_{obs}$  is the number of observed TMHs, and  $N_{prd}$  is the total number of predicted TMHs. However, this statistical measure needs the availability in advance of the real (observed) helices, which is not the case in our work. Nevertheless, this has been applied in Zemni et al. (2019a), where the present method has been applied to a well-known example where the observed TMHs are known, and leads there to a  $Q_p = 100\%$ .

Another statistical measure is applied in Ben Mabrouk and Ibrahim Mahmoud (2013), Ben Mabrouk et al. (2015), Bin and Zhang (2013), Ibrahim Mahmoud et al. (2016) and Zemni et al. (2019a) consisting in a prediction certainty evaluation based on the computation of a type of an absolute deviation via the difference between the first residue observed and the first predicted residue. More precisely, we call mean absolute error (MAE) the quantity

$$MAE = \sum_{i=1}^K |a_i^{obs} - a_i^{prd}| + |b_i^{obs} - b_i^{prd}|,$$

where  $[a_i^{obs}, b_i^{obs}]$ , are the observed (real) TMHs segments,  $[a_i^{prd}, b_i^{prd}]$ , are the predicted TMHs segments, and  $K$  the total number of segments (real or observed). The MAE should be as small as possible. Remark that this measure requires also the real segments to be available.

Nevertheless, in the present paper, we proposed to check the accuracy of our method by applying a different statistical measure based on the so-called Jack-Knife test (Calvo et al. 2005; Rezaei et al. 2008) according to the same criterion used in the previous existing works consisting in the number of the segments predicted, or equivalently to the number of maximum peaks.

The basic idea of the Jack-knife tests consists in recursively deleting a single observation from the sample, and compute the estimation until there are  $n$  estimates for a sample size of  $n$ . In our case, we will consider a modified version of such a test, in which, we consider for each level  $J$  the data points  $N_J$  designating the number of predicted peaks (segments) for the approximation component  $A_J$  of the numerical series associated to the protein strain. We thus compute the estimator  $J$  times, for:

- $N_i, 1 \leq i \leq J,$
- $N_i, 2 \leq i \leq J,$
- $\dots,$
- $N_{J-1}, N_J,$
- $N_J.$

Once, the  $J$  estimates  $\widehat{N}_1, \widehat{N}_2, \dots, \widehat{N}_J$ , are obtained, the standard error is calculated as

$$Se_{Jackknife}(J) = \sqrt{\frac{J-1}{J} \sum_{i=1}^J (\widehat{N}_i - \widehat{N}_0)^2},$$

where  $\widehat{N}_0$  is the arithmetic mean of the vector  $(\widehat{N}_i)_{1 \leq i \leq J}$ . Such a test is estimated approximately to 95.6% reflecting a good performance of the method, and thus a good localization of the desired segments.

Notice that the example studied here is an important case that may be considered as a model to be applied to the new case of the coronavirus COVID-19 when a database is available which is not the case for us. We also mention that the wavelet/multi-wavelet theory are proved to be effective in discovering, and identifying abnormalities, and special facts in biological strings such as helices, knots,.... Thus, with no laboratory study available on the chain used here, and its equivalents in the new COVID-19, we intend that the current study may be applied to identify such abnormalities, and other characteristics for the new virus COVID-19 chains as well as other cases. A step forward in the application of the present method has been conducted in Zemni et al. (2019a) where the statistical measures evoked above has been applied

on a more detailed example of biological series with observed segments already available. This permitted to compute the index  $Q_p$  and the error MAE discussed above.

Transmembrane proteins are generally composed of more than 18 amino acids, sometimes 30. Therefore, our method is mainly to detect the approximate locations of these segments around the spikes. This will be important for practitioners as it tells them the approximate location of the anomaly, if any, before starting any experimental trial. This will be confirmed by comparison with the segments observed experimentally when possible, as in Zemni et al. (2019a) for the case of HSch entropy measure, or also (Ben Mabrouk and Ibrahim Mahmoud 2013; Ben Mabrouk et al. 2015; Fischer et al. 2003; Ibrahim Mahmoud et al. 2016) in the case of single wavelets.

## 5 Conclusion

In this paper, multi-wavelet procedure has been developed extending the well known wavelet algorithms applied in image, and signal analysis. By improving the existing ideas on multi-wavelets, we constructed new ones, and proved that multi-filters may be associated, and applied in signal analysis with more efficient results compared to the classical ones. Error estimates as well as fast algorithms have been proved, and applied on ECG signals, and a coronavirus case. Among the theoretical findings of the paper which may be resumed in the construction of the HSch multi-wavelet, its filters, as well as their matrix representation, the experimental findings may be resumed in three directions. In a first experimentation, an approximation (reconstruction) of a classical example dealing with Fourier modes has been conducted. Such an example may be seen as a universal model for periodic and stationary signals which are generally well approximated (reconstructed) even with classical methods such as Fourier one. The present methods has lead to suitable error estimates as well as fast algorithms. The second experimentation has been concerned with the HSch multi-wavelet de-noising and reconstruction of a benchmark (highly volatile, non stationary) signal due to an ECG case. Our method has been proved here also to be efficient in approximating such signals. The last experimentation is concerned with a denoising case applied on a strain of coronavirus signal due to the Cov-2 (SARS). The idea has turned around the localization of the transmembrane segments of such a series as local maxima of an numerized version of the strain obtained by Kyte–Doolittle method (Kyte and Doolittle 1982). Accuracy of the method has been evaluated by means of error estimates and statistical tests. The present method has been already applied on more examples where more statistical tests and errors may be applied (Zemni et al. 2019a). We intend finally to continue exploring wavelet and multi-wavelet method in investigating

MPIPLLTLTSGSDLRCTTPDDVQAPNYTQHTSSMRGVVYPDEIPRSDTLYLTQDLPPLPYSNVTGPHTINHPTGNPVIPPKD  
GIYPAATEKSNVVRGWVPGSTMNNKSQSVIIINNSTNVVIRACNPELCDNPPPAVSKPMGTQTHTMIPDNAPNCTPEYISDAPS  
LDVSEKSGNPKHLREPVPKNKDGPLYVYKGYQPIDVVRDLPSGPNLTKPIKPLGINITNPRAILTAPSPAQDIWGTSAAYVVG  
YLKPTTPMLKYDENGITDAVDCSQNPLAELKCSVKSPKIDKGIYQTSNPRVVPVSGDVRPPNITNLCPPEVNPATKPPSVYAW  
ERKKISNCVADYSVLYNSTPSTPKCYGVSATKLNLDLCPNSVYADSPVVKGDDVVRQIAPGQTGVIADYNYKLPDDPMGCVLAW  
NTRNIDATSTGNYYKYRYLRHGKLRPPERDISNVPPSPDGKCTPPALNCYWPLNDYGPYTTTGIGYQPYRVVLSPELLNAPA  
TVCGPKLSTDLIKDQCVNPNPNGLTGTGVLTPSSKRPPQPPQGRDVSPTDSVRDPKTSEILDISPCSPGGVSVITPGTNASSEV  
AVLYQDVNCTDVSTAIHADQLTPAWRIYSTGNNVPTQAGCLIGAEHVDTSECDIPIGAGICASYHTVSLRSTSQKSIVAYTM  
SLGADSSIAYSNNTIAIPTNPSISITTEVMPVSMAKTSVDCNMYICGDSTECANLLQYGSPTQLNRALSGLIAAEQDRNTREVA  
QVKQMYKTPTLKYGGPNPSQILPPDPLKPTKRSPIEDLLPNKVTLADAGPMKQYECLGDINARDLCAQKPNGLTVLPLLTDD  
MIAAYTAALVSGTATAGWTPGAGAALQIPPAMQAMAYRPNIGVGTQNVLYENQKQIANQPNKAISQIQESLTTTSTALGKLDQ  
VVNQNAQALNTLVKQLSSNPGAISSVLNDILSRDKVEAEVQIDRLITGRQLSQTYYVTTQLIRAAEIRASANLAATKMSECVLGG  
SKRVDPCGKGYHLMSPQAAPHGVVPLHVTVYVPSQERNPTTAPAICHEGKAYPPREGVVPVNGTSSWPITHQRNPPSPQIITD  
NTPVSGNCDVVIGIINNTVYDPLQPELDSPEELDKYKPNHTSPDVLGDISGINASVUNIQKEIDRLNEVAKNLNLSLIDLQELGK  
YEQYIKWPWYVWLGPIAGLIAIVMVTILLCCMTSCCSCLKGACSCGSCCKPDEDDSEPVLKGVKLYHT

**Fig. 12** The coronavirus proteins' series strain

complex cases such as COVID-19 cases by testing many cases of wavelets multi-wavelets and more statistical tests, and also more complicated cases of molecular/cellular communications signals. Recall that, even from the theoretical point of view, the choice of the model, the estimating, and/or the analyzing bases is always and already the most hard task.

**Acknowledgements** The authors would like to recall that the present work in its original version has been devoted to the theoretical framework of the multi-wavelets processing. An application part of the work has been the object of a conference talk (Zemni et al. 2019b). Some extending results based on the present paper have also been the subject of another paper (Zemni et al. 2019a), although based on the present one, but has been published before it due to the long process in the publication of the present one.

**Funding** The authors declare that no funds or grants have been received from any direction.

## Declarations

**Conflict of interest** The authors declare that no conflict of interest for the present work.

**Ethical approval** The authors declare that no animals were involved in the study, and that this article does not contain any studies with human participants performed by any of the authors.

## 6 Appendix: The coronavirus proteins' series strain

We found difficulties in obtaining the strain protein in hand. For this reason, we provided it in this appendix manually to be in the disposition of readers, and researchers (Fig. 12).

## References

- Abhijith A, Ruban DP, Rajy X, Mareeta CP (2016) Review of signal processing techniques for detection of power quality events. *Am J Eng Appl Sci* 9(2):364–370
- AlMahamdy M, Riley HB (2014) Performance study of different denoising methods for ECG signals. *Procedia Comput Sci* 37:325–332
- Alramahi N, Bush M, Swash MR (2018) Numerically analysed multiwavelet transform computations: multidimensional compression case studies. *J Fundam Appl Sci*. 10(4):691–693
- Alwan IMG (2014) Image steganography by using multiwavelet transform. *Baghdad Sci J* 11(2):275–283
- Anand K, Palm GJ, Mesters JR, Siddell SG, Ziebuhr J, Hilgenfeld R (2002) Structure of coronavirus main proteinase reveals combination of a chymotrypsin fold with an extra  $\alpha$ -helical domain. *EMBO J* 21(13):3213–3324
- Arfaoui S, Rezgui I, Ben Mabrouk A (2017) Wavelet analysis on the sphere, spheroidal wavelets. Degryuter. ISBN 978-3-11-048188-4
- Arfaoui S, Ben Mabrouk A, Cattani C (2020a) New type of Gegenbauer–Hermite monogenic polynomials, and associated continuous Clifford wavelet transform. *J Math Imaging Vis* 62:73–97. <https://doi.org/10.1007/s10851-019-00914-y>
- Arfaoui S, Ben Mabrouk A, Cattani C (2020b) New type of Gegenbauer–Jacobi–Hermite monogenic polynomials, and associated continuous Clifford wavelet transform. *Acta Applicandae Mathematicae* 170(1):1–35. <https://doi.org/10.1007/s10440-020-00322-0>
- Attakitmongkol K, Hardin DP, Wilkes DM (2001) Multiwavelet prefilters—part II: optimal orthogonal prefilters. *IEEE Trans. Image Process* 10(10):1476–1487
- Babatunde SE (2012) A review of signal processing techniques for heart sound analysis in clinical diagnosis. *J Med Eng Technol* 36(6):303–307
- Bacchelli S, Cotronei M, Sauer T (2002) Multifilters with, and without prefilters. *BIT* 42:231–261
- Ben Mabrouk A, Kortas H, Dhifaoui Z (2008a) A wavelet support vector machine coupled method for time series prediction. *Int J Wavelets Multiresolution Inf Process* 6(6):1–17
- Ben Mabrouk A, Ben Abdallah N, Dhifaoui Z (2008b) Wavelet decomposition, and autoregressive model for the prevision of time series. *Appl Math Comput* 199(1):334–340
- Ben Mabrouk A, Ibrahim Mahmoud MM (2013) Multifractal study of some biological series, the case of proteins. *IWSSIP*

- Ben Mabrouk A, Rabbouch B, Saadaoui F (2015) A wavelet based methodology for predicting transmembrane segments. In: Poster session, the international conference of engineering sciences for biology, and medicine, 1–3 May 2015, Monastir, Tunisia
- Bin Y, Zhang Y (2013) A simple method for predicting transmembrane proteins based on wavelet transform. *Int J Biol Sci* 9(1):22–33
- Bonnin A (2018) Caractérisation de la protéine S du coronavirus humain 229E. These do Doctorat de l'Université de Lille, Ecole Doctorale Biologie-Santé de Lille
- Brazile CR (2009) B.S, multivariate multiresolution multiwavelets, Thesis in mathematics. Texas Tech University
- Bui TD, Chen G (1998) Translation-invariant denoising using multiwavelets. *IEEE Trans Signal Process* 46:3414–3420
- Calvo JA, Millett KC, Rawdon EJ, Stasiak A (2005) Physical and numerical models in knot theory including applications to the life sciences. In: *Series on knots and everything*, vol 36. World Scientific, Singapore
- Cotronei M, Montefusco LB, Puccio L (1998) Multiwavelet analysis and signal processing. *IEEE Trans Circuits Syst II Analog Digital Signal Process* 45(8):970–987
- Cotronei M, Puccio L (1997) An application of multiwavelet analysis to signal compression. In: Le Mehaute A, Rabut C, Schumaker L (eds) *Surface fitting and multiresolution methods*. Vanderbilt Univ. Press, Nashville, pp 75–82
- Cotronei M, Sissouno N (2017) A note on Hermite multiwavelets with polynomial, and exponential vanishing moments. 17 pages [arXiv:1702.01007v1](https://arxiv.org/abs/1702.01007v1) [math.NA] (2017)
- Daubechies I (1992) Ten lectures on wavelets. SIAM Philadelphia
- Desjardins J (2010) Roles des protéines accessoires du coronavirus humain OC43. Institut Armand Frappier, Université du Québec, Mémoire Maître es-Sciences, INRS
- Dauvergne J (2010) Synthèse et étude physico-chimique de nouveaux tensioactifs utilisables pour la cristallisation 2D sur film lipidique et l'étude des protéines membranaires. Thèse de doctorat de l'Université d'Avignon et des Pays du Vaucluse
- Efromovich S (2001) Multiwavelets, and signal denoising. *Sankhya Ser A* 63:367–393
- Fischer P, Baudoux G, Wouters J (2003) Wavepred: a wavelet-based algorithm for the prediction of transmembrane proteins. *Commun Math Sci* 1(1):44–56
- Fowler JE, Hua L (2002) Wavelet transforms for vector fields using omnidirectionally balanced multiwavelets. *IEEE Trans Signal Process* 50(12):3018–3027
- Geronimo JS, Hardin DP, Massopust PR (1994) Fractal function and wavelet expansions based on several scaling functions. *J Approx Theory* 78:373–401
- Hardin DP, Roach DW (1998) Multiwavelet prefilter-I: orthogonal prefilter preserving approximation order  $p \leq 2$ . *IEEE Trans Circuits Syst II Analog Digit Signal Process* 45(8):1106–1112
- Ho YF (2002) Singularity detection for regularity scalable image coding. Master thesis, Department of electronic and information engineering, Hong Kong University of Technology
- Ho CYF, Ling WK, Tam KS (2003) Denoising by multiwavelet singularity detection. In: *Proceedings of the 2003 International conference on neural networks and signal processing*, December 2003, Nanjing, vol 1, pp 616–619
- Huang Ch-P, Li Ch-Ch (2011) Secret image sharing using multiwavelet transform. *J Inf Sci Eng* 27:733–748
- Ibrahim Mahmoud MM, Ben Mabrouk A, Hashim MA (2016) Wavelet multifractal models for transmembrane proteins' series. *Int J Wavelets Multires Inf Process* 14(6):36
- Iyer LR (2001) Image compression using balanced multiwavelets. Master thesis of Science in Electrical Engineering, Faculty of the Virginia Polytechnic Institute, and State University
- Jallouli M, Zemni M, Ben Mabrouk A, Mahjoub MA (2019a) Toward recursive spherical harmonics-issued bi-filters: part I: theoretical framework. *Soft Comput* 23:10415–10428. <https://doi.org/10.1007/s00500-018-3596-9>
- Jallouli M, Belhadj Khalifa W, Ben Mabrouk A, Mahjoub MA (2019b) Toward recursive spherical harmonics issued bi-filters: Part II: an associated spherical harmonics entropy for optimal modelling. *Soft Comput* 1–13. <https://doi.org/10.1007/s00500-019-04274-y>
- Jallouli M, Belhadj Khalifa W, Ben Mabrouk A, Mahjoub MA (2019c) Toward new spherical harmonic shannon entropy for surface modeling. In: Vento M, Percannella G (eds) *Computer analysis of images, and patterns*. CAIP 2019. Lecture Notes in Computer Science, vol 11679. Springer, Cham, pp 38–48
- Jallouli M, Belhadj Khalifa W, Ben Mabrouk A, Mahjoub MA (2019d) Assessment of global left ventricle deformation using recursive spherical harmonics. In: Liu Y, Wang L, Zhao L, Yu Z (eds) *Advances in natural computation, fuzzy systems, and knowledge discovery*. ICNC-FSKD 2019. Advances in intelligent systems, and computing, vol 1074. Springer, Cham. <https://doi.org/10.1007/978-3-030-32456-8-54>
- Johnson BR (2000) Multiwavelet moments, and projection prefilter. *IEEE Trans Signal Process* 48:3100–3108
- Keinert F (2004) Wavelets, and multiwavelets. Chapman & Hall/CRC
- Kessler B (2009) Multiwavelets for quantitative pattern matching. In: *Proceedings of the 42nd Hawaii international conference on system sciences—2009*, 10 pages
- Kotas M, Moron T (2017) ECG signals reconstruction in subbands for noise suppression. *Biocybern Biomed Eng* 37(3):453–465
- Kyte J, Doolittle RF (1982) A simple method for displaying the hydrophathic character of a protein. *J Mol Biol* 157:105–132
- Lebrun J, Vetterli M (1998) Balanced multiwavelets theory, and design. *IEEE Trans Signal Process* 46:1119–1125
- Lebrun J, Vetterli M (2001) High-order balanced multiwavelets: theory, factorization, and design. *IEEE Trans Signal Process* 49:1918–1930
- Li F (2016) Structure, function, and evolution of coronavirus spike proteins. *Annu Rev Virol*. 3(1):237–261. <https://doi.org/10.1146/annurev-virology-110615-042301>
- Liang K-C, Li J, Jay Kuo C-C (1996) Image compression with embedded multiwavelet coding. In: *Proceedings of SPIE—the international society for optical engineering*, 12 pages. <https://doi.org/10.1117/12.235990>
- Liu M, Gu Ch, Wu J, Zhu Y (2006) Amino acids 1 to 422 of the spike protein of SARS associated coronavirus are required for induction of cyclooxygenase-2. *Virus Genes* 33:309–317. <https://doi.org/10.1007/s11262-005-0070-4>
- Mallat S (2008) A wavelet tour of signal processing, 3rd edn. Academic Press, Cambridge
- Massopust PR, Ruch DK, Van Fleet PJ (1996) On the support properties of scaling vectors. *Appl Comput Harmon Anal* 3:229–238
- McBride R, Van Zyl M, Fielding BC (2014) The coronavirus nucleocapsid is a multifunctional protein. *Viruses* 6:2991–3018. <https://doi.org/10.3390/v6082991>
- Millett KC (2005) Tying down open knots: a statistical method for identifying open knots with applications to proteins. In: Calvo JA, Millett KC, Rawdon EJ, Stasiak A (eds) Chapter 11 in *physical and numerical models in knot theory including applications to the life sciences*. Series on knots and everything, vol 36. World Scientific, Singapore, pp 203–217
- Rezaei MA, Abdolmaleki P, Karami Z, Asadabadi EB, Sherafat MA, Abrishami-Moghaddam H, Fadaie M, Forouzanfar M (2008) Prediction of membrane protein types by means of wavelet analysis and cascaded neural networks. *J Theor Biol* 254:817–820
- Rieder P, Götze J, Nossek JA (1996) multiwavelet transforms based on several scaling functions. In: *Proceedings of IEEE-SP international symposium on time-frequency, and time-scale analysis*, pp 309–312

- Ruedin AMC (2002) Construction of nonseparable multiwavelets for nonlinear image compression. *Eurasip J Appl Signal Process* 1:73–79
- Ruedin AMC (2003) Nonseparable multiwavelets: construction, and applications to image processing. PhD thesis, Universidad de Buenos Aires
- Selesnick IW (1998) Multiwavelet bases with extra approximation properties. *IEEE Trans Signal Process* 46:2898–2908
- Selesnick IW (1999) Interpolating multiwavelet bases, and the sampling theorem. *IEEE Trans Signal Process* 47:1615–1621
- Selesnick IW (2000) Balanced multiwavelet bases based on symmetric FIR filters. *IEEE Trans Signal Process* 48:184–191
- Shen L, Tan HH (2001) On a family of orthonormal scalar wavelets, and related balanced multiwavelets. *IEEE Trans Signal Process* 49:1447–1453
- Stacey AJ, Blyth WF (2008) A variational method using Alpert multiwavelets. *ANZIAM J* 48:820–836
- Stankovic RS, Falkowski BJ (2003) The Haar wavelet transform: its status, and achievements. *Comput Electr Eng* 29:25–44
- Strela V, Heller PN, Strang G, Topiwala P, Heil C (1999) The application of multiwavelet filter banks to image processing. *IEEE Trans Signal Process* 8:548–563
- Talbot P, Jouvenne P (1992) Le potentiel neurotrope des coronavirus. *médecine/sciences* 8(2):119–125
- Tham JY, Shen L, Lee SL, Tan HH (2000) A general approach for analysis, and application of discrete multiwavelet transform. *IEEE Trans Signal Process* 48:457–464
- Turcajova R (1999) Construction of symmetric biorthogonal multiwavelets by lifting, in *Wavelet applications in signal, and image processing VII*. In: Unser MA, Aldroubi A, Laine AF (eds) SPIE, vol 3813, pp 443–454
- Van Der Werf S (2010) Nouvelle souche de coronavirus associée au SRAS et ses applications. EP1 694 829 B1, Fascicule de brevet européen
- Vehel MJ, Aldroubi A (1997) Projection based pre-filtering for initialising multiwavelet algorithms. In: *Proc. SPIE wavelet applications in signal and image processing*, vol 3169, pp 432–437
- Wang Z, Wan F, Wong CM, Zhang L (2016) Adaptive Fourier decomposition based ECG denoising. *Comput Biol Med* 77(1):195–205
- Xia X-G (1998) A new prefilter design for discrete multiwavelet transforms. *IEEE Trans Signal Process* 46:1558–1570
- Xia T, Jiang Q (1999) Optimal multifilter banks: design, related symmetric extension transform, and application to image compression. *IEEE Trans Signal Process* 47:1878–1889
- Xia X-G, Suter BW (1996) Vector-valued wavelets and vector filter banks. *IEEE Trans Signal Process* 44(3):508–518
- Xia X-G, Geronimo JS, Hardin DP, Suter BW (1996) Design of prefilters for discrete multiwavelet transforms. *IEEE Trans Signal Process* 44:25–35
- Xu L, Zhang F, Yang W, Jiang T, Lu G, He B, Li X, Hu T, Chen G, Feng Y, Zhang Y, Fan Q, Feng J, Zhang H, Tu Ch (2016) Detection, and characterization of diverse alpha-, and beta coronaviruses from bats in China. *VIROLOGICA SINICA* 31(1):69–77. <https://doi.org/10.1007/s12250-016-3727-3>
- Yoganand S, Mohan BM (2018) Denoising of ECG signals using multiwavelet transform. *Helix* 8(1):2696–2700
- Zemni M, Jallouli M, Ben Mabrouk A, Mahjoub MA (2019a) Explicit Haar–Schauder multiwavelet filters, and algorithms. Part II: relative entropy-based estimation for optimal modeling of biomedical signals. *Int J Wavelets Multiresolution Inf Process* 17(05):1950038
- Zemni M, Jallouli M, Ben Mabrouk A, Mahjoub MA (2019b) ECG signal processing with Haar–Schauder multiwavelet. In: *Proceedings of the 9th international conference on information systems, and technologies—ICIST*. <https://doi.org/10.1145/3361570.3361611>
- Zhang J-K, Davidson TN, Luo Z-Q, Wong KM (2001) Design of interpolating biorthogonal multiwavelet systems with compact support. *Appl Comput Harmon Anal* 11:420–438

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.