Diagonal time dependent state space models for modal decomposition of non–stationary signals[☆]

Luis Enrique Avendaño[1], Luis David Avendaño-Valencia[2,☆☆], and Edilson Delgado Trejos[3]

[1] Escuela de Tecnología Eléctrica, Universidad Tecnológica de Pereira, Pereira, Colombia.

[2] Institute for Structural Engineering, Department of Civil and Geomatic Engineering, ETH Zürich. HIL H33.1, Stefano-Franscini-Platz 5, Zürich, Switzerland.

[3] Quality and Production Department, Instituto Tecnológico Metropolitano de Medellín, (ITM). Calle 73 No. 76A-354. Medellín, Colombia.

Abstract

This work is devoted to the problem of the decomposition of a non-stationary signal into modal components, for which a methodological approach based on diagonal time-dependent state space models is postulated. In particular, on this paper is shown that the response of a diagonal time-dependent state space models can be cast into a modal form characterized by time-dependent amplitudes and frequencies. Later, a Kalman filter based framework for non-stationary modal decomposition is built on the previously discussed diagonal state space representations. The enhanced performance of the proposed methods is demonstrated on a benchmark test consisting of three non-stationary modal components, and on the modal decomposition and denoising of a *ElectroCardio Graphic* signals from the QT database. The proposed methods constitute a reliable tool for on-line modal decomposition of multi-component non-stationary signals, with results comparable and even better than other state-of-the-art methods.

Keywords: Linear time-dependent state space, Kalman filtering, non-stationary modal decomposition.

1. Introduction

Most real-life systems are characterized by time-dependent dynamics, while their resulting dynamic response is non–stationary. The effective analysis of non-stationary signals requires appropriate representation methods capable of accurately describing

[☆]**Please cite as:** Luis Enrique Avendaño, Luis David Avendaño-Valencia, Edilson Delgado-Trejos, *"Diagonal time dependent state space models for modal decomposition of non-stationary signals*, Signal Processing, Volume 147, June 2018, Pages 208-223, https://doi.org/10.1016/j.sigpro.2018.01.031.

[☆☆]Corresponding author. E-mail: avendano@ibk.baug.ethz.ch Tel: (+41) 44 633 79 95.

the evolutionary signal dynamics [1, Ch. 5]; [2, Ch. 1]; [3]. Amongst the most recognizable representation methods of non–stationary signals are non–parametric Time–Frequency Representations (TFR) and Time–Scale Representations (TSR), where the non–stationary signal is characterized in terms of an infinite set of oscillatory components localized in both time and frequency [4, Ch. 2]; [1, Ch. 5]. Modal representations, on the other hand, aim at representing a non–stationary signal by means of the superposition of a finite number of oscillatory components, referred to as *modes*. Unlike the stationary case where the amplitude and frequency of each mode are deemed constant, in the non-stationary case, each mode is associated with time-dependent amplitude and frequency values [5]. In contrast to TFR and TSR where representation is hinged on surfaces on the time-frequency plane, modal representations offer more compact representations of non-stationary processes, while facilitate the extraction of individual signal components.

The problem of calculating the modal representation of a given signal shall hereby be referred to as *modal decomposition*. In a modal representation, the signal may consist of a single modal component (thus called mono-component) or may be constructed from the superposition of several modal components (thus referred to as multi-component). The calculation of the modal decomposition of a mono-component signal reduces to the calculation of the *instantaneous frequency* (IF), *instantaneous phase* (IP) and *instantaneous amplitude* (IA) of the signal [6, 7, 8]. In addition, modal decomposition in the case of multi-component signals further requires the isolation of individual modal components and the estimation of their respective instantaneous modal parameters [7, 8]. Main difficulties associated with the decomposition of multiple modal components stem from the precise localization and tracking of each evolutionary frequency trajectory in a noisy environment, especially in the case of components with crossing frequencies or in the case of vanishing components.

Among the most widely appraised modal decomposition methods are those based on ridge extraction from TFRs or TSRs. Indeed, for noise-free mono-component signals the IF is located at the maximum of the TFR while the IA is associated with the value at the ridge [9]. Then, estimation of the IF of the modal component translates into the localization of the maximum in a TFR. Nonetheless, in noisy signals, IF estimates are noisy as well and often exhibit discontinuities. To this end, more specialized methods based on curve extraction procedures, such as the Viterbi algorithm, have been developed [6, 8]. In multi-component signals, TFR-based modal decomposition may be achieved by the *peeling* method, where the most dominant modal component is extracted from the TFR and subsequently removed from the signal to extract other modal components [7, 8, 10]. Otherwise image processing algorithms are also available to extract and separate ridges on TFR images [11]. Nonetheless, these specialized TFR-based modal decomposition methods require heavy processing on time-frequency surfaces, while their performance is bound to the method of calculation of the TFR and a careful parameter selection.

Other modal decomposition methods are based on different types of non–parametric non–stationary representations of the signal. For instance, the *Discrete Wavelet Transform* (DWT) and related filter–bank based methods attempt at separating the signal into frequency bands [12, 13]. Otherwise, the *Empirical Mode Decomposition* or *Hilbert–Huang Transform* aims at separating the signal into various non–stationary orthogonal

signal components referred to as *Intrinsic Mode Functions* from which IA and IF estimates may be derived through the use of the Hilbert transform [5, 14, 15, 16, 13]. Other methods achieve a modal decomposition by eigen-analysis of the signal autocovariance matrix. Among these, *Subspace Methods* and the *Karhounen–Loève Transform* (also referred to as Principal Component Analysis, Proper Orthogonal Decomposition or Singular Spectrum Analysis) are available [17, 18, 19, 20]. A common problem of these methods is that often the resulting decomposition is characterized by modes that have no relation with physical features of the underlying system. Additionally, these methods cannot resolve effectively the cases of crossing or vanishing modes.

Alternatively, it is possible to define a certain parametric structure associated with a modal form of the signal of interest. A simple example is formed by adaptive notch filters, whose central frequency and bandwidth is adaptively adjusted to the signal properties, while their performance is defined by a forgetting factor, which sets a trade–off between the tracking accuracy and the estimation error [21, 22, 23]. More powerful methods attempt at tracking the IF and IA of a single modal component either by adaptive tracking [24, 25], or by projecting their values into a functional basis [26]. The latter methods are hinged on non-linear optimization methodologies, and as a result their performance is driven by a correct choice of initial conditions to ensure convergence to a global maximum. In turn, the presence of noise will significantly affect the overall performance of these methods.

Linear time-dependent state space representations and their properties may be used with the purpose of calculating modal decompositions of non-stationary signals. One of such methods is based on a particular type of block-diagonal state space representation, where each second–order block is associated with a modal component with specific IA and IF. In turn, each block is parametrized, through a non-linear relation, by the IF of the modal component. Then, a modal decomposition is achieved by joint estimation of the state vector and the instantaneous frequency by means of the Extended Kalman Filter (EKF) or other non-linear state estimation methods. These methods are commonly referred to as *Kalman Filter Frequency Trackers* (KF-FT) [27, 28, 29, 30, 31, 32]. Similarly, an equivalence transformation may be used to map the state space representation associated with a *Time-dependent AutoRegressive* (TAR) model of the signal of interest, into a diagonal state space representation, where each entry of the state vector of the diagonalized system is associated with a modal component, while the entries in the diagonal of the state matrix are associated with the IF of the respective modal components [33, 34]. After having selected the number of modal components, state-space based modal decomposition methods can lead to on-line estimates of the IF, IA and the modal components on multi-component signals. In addition, these methods can be modified for the case of vanishing components, by on-line adjustment of the number of modal components.

This work is a further contribution towards the class of modal decomposition methods based on linear time-dependent state-space models. In particular, the main aim of this paper is to study the properties of the response of linear time-dependent state-space systems and to demonstrate how such a response can be associated with a modal representation by the use of equivalence transforms. In that sense, and in contrast to most recent studies aiming at describing the response of a linear time-dependent system in terms of Instantaneous and Harmonic Frequency Response Functions [35, 36], the

main contribution of this work is to show, through the use of equivalence transforms, that the response of a linear time-dependent state-space system may be characterized by the superposition of narrowband frequency components, which in the case of the natural response take the form of modal components with IA and IF associated with the instantaneous eigenvalues of second-order time-dependent state-space blocks. Furthermore, the analysis undertaken in this work leads to a more general and yet simpler parametrization of second-order linear time-dependent state space blocks, which then can be used in combination with Kalman filters to yield accurate on-line estimates of multiple modal components and their respective IA and IF. Two estimation approaches are considered based either on Joint or Decoupled Kalman filter methods, which lead to the proposed *Joint/Dual Kalman Filter Non-stationary Sinusoid Tracking* (JKF-NST or DKF-NST) methods. The postulated JKF-NST and DKF-NST methods turn out to be a generalization of the well-known KF-FT method, with the main advantage being that the concurrent state and parameter estimation problem involves a milder type of non-linearity compared to that one appearing in KF-FTs.

A Monte Carlo analysis involving a three–component non–stationary signal with mode crossing and fading under noise, and the application on the denoising of Electro-Cardio Graphic (ECG) records from the QT database, demonstrate the increased robustness of the proposed methods compared to the KF-FT method and the ridge extraction method described in [8], in terms of tracking accuracy of the modal components and their IA and IF. Moreover, the theoretical results obtained from the analysis of linear time-dependent state space models may be further extended towards the design of improved modal decomposition methods.

The paper is organized as follows: Section 2 provides the main terminology and definitions used throughout the work, Section 3 provides the main contributions of this work, where it is shown how the response of a complex diagonal state space system may be cast into a modal form, Section 4 introduces the second order block diagonal structures to be later used in Section 5 for estimation of modal decompositions via Kalman filter methods. Section 6 and Section 7 show a performance analysis of the modal decomposition methods in the decomposition of a three component non-stationary signal featuring mode crossing and fading, and in the denoising of ECG records. Finally, the conclusions of this work are summarized in Section 8.

2. Time-dependent state space representations - Main definitions

Consider the (discrete-time) time-dependent state space representation:

$$\boldsymbol{z}[t] = \boldsymbol{A}[t] \cdot \boldsymbol{z}[t-1] + \boldsymbol{b}[t] \cdot u[t] \tag{1a}$$

$$y[t] = \boldsymbol{c}[t] \cdot \boldsymbol{z}[t] + w[t], \qquad w[t] \sim \texttt{NID}\left(0, \sigma_w^2\right) \tag{1b}$$

where $t \in \mathbb{Z}^+$ is the normalized discrete time, $y[t] \in \mathbb{R}$ is the response signal, $\boldsymbol{z}[t] \in \mathbb{R}^M$ is the *state vector*, $u[t] \in \mathbb{R}$ is the *excitation* of the system, while $\boldsymbol{A}[t] \in \mathbb{R}^{M \times M}$, $\boldsymbol{b}[t] \in \mathbb{R}^{M \times 1}$, and $\boldsymbol{c}[t] \in \mathbb{R}^{1 \times M}$ are the *time-dependent state matrix*, *input coupling vector*, and *state measurement vector*. In addition, $w[t]$ is a *Normally and Identically Distributed* (NID) innovations with zero mean and variance σ_w^2.

The state of the system and its response at time t may be computed based on an initial state $\boldsymbol{z}[0]$ and excitation $u[t]$, by means of the expressions [37, pp. 110–111]:

$$\boldsymbol{z}[t] = \boldsymbol{\Phi}[t,0]\cdot\boldsymbol{z}[0] + \sum_{\tau=1}^{t}\boldsymbol{\Phi}[t,\tau]\cdot\boldsymbol{b}[\tau]\cdot u[\tau] \tag{2a}$$

$$y[t] = \boldsymbol{c}[t]\cdot\boldsymbol{\Phi}[t,0]\cdot\boldsymbol{z}[0] + \sum_{\tau=1}^{t}\boldsymbol{c}[t]\cdot\boldsymbol{\Phi}[t,\tau]\cdot\boldsymbol{b}[\tau]\cdot u[\tau] + w[t] \tag{2b}$$

where $\boldsymbol{\Phi}[t,\tau] \triangleq \boldsymbol{A}[t]\cdot\boldsymbol{A}[t-1]\cdots\boldsymbol{A}[\tau], \forall t > \tau$ is referred to as the *state transition matrix*. Note that the first component on the right hand side of Equation (2b) corresponds to the *natural response* of the system, while the second component accounts for the *forced response*, respectively.

The state space representation in Equation (1) may be associated with an *equivalent* state space representation by means of an equivalence transformation. For that purpose, let $\boldsymbol{V}[t] \in \mathbb{R}^{M\times M}$ be a non-singular orthogonal matrix satisfying $\boldsymbol{V}[t]\cdot\boldsymbol{V}^{-1}[t] = \boldsymbol{I}$, and $\bar{\boldsymbol{z}}[t] = \boldsymbol{V}[t]\cdot\boldsymbol{z}[t]$. Then, the state space equation:

$$\bar{\boldsymbol{z}}[t] = \bar{\boldsymbol{A}}[t]\cdot\bar{\boldsymbol{z}}[t-1] + \bar{\boldsymbol{b}}[t]\cdot u[t] \tag{3a}$$

$$y[t] = \bar{\boldsymbol{c}}[t]\cdot\bar{\boldsymbol{z}}[t] + w[t] \tag{3b}$$

where:

$$\bar{\boldsymbol{A}}[t] = \boldsymbol{V}[t]\cdot\boldsymbol{A}[t]\cdot\boldsymbol{V}^{-1}[t-1], \qquad \bar{\boldsymbol{b}}[t] = \boldsymbol{V}[t]\cdot\boldsymbol{b}[t], \qquad \bar{\boldsymbol{c}}[t] = \boldsymbol{c}[t]\cdot\boldsymbol{V}^{-1}[t] \tag{4}$$

is said to be *(algebraically) equivalent* to Equation (1), and the matrix $\boldsymbol{V}[t]$ is called an *equivalence* transformation [38]; [39]; [37, pp. 111–115].

In the same form, the state transition matrix $\boldsymbol{\Phi}[t,\tau]$ associated with the state space representation in Equation (1) is transformed according to the expression:

$$\bar{\boldsymbol{\Phi}}[t,\tau] = \boldsymbol{V}[t]\cdot\boldsymbol{\Phi}[t,\tau]\cdot\boldsymbol{V}^{-1}[\tau-1] \tag{5}$$

where $\bar{\boldsymbol{\Phi}}[t,\tau]$ is the *equivalent* state transition matrix.

A consequence of the above is that although the state transition matrices are different, the response of the two equivalent systems is exactly the same. This property applies even on the time-dependent case and can be used to facilitate the representation or the evaluation of the response of a system described by a state space representation. Nonetheless, an important difficulty on the time-dependent case stems from the fact that in general the transformation matrix is also time-dependent.

3. Modal solution of a diagonal time-dependent state space representation

Consider a time-dependent system characterized by the (complex) diagonal time-dependent state matrix:

$$\boldsymbol{A}[t] \triangleq \boldsymbol{\Lambda}[t] = \operatorname{diag}(\boldsymbol{\lambda}[t]) = \begin{bmatrix} \lambda_1[t] & & \\ & \ddots & \\ & & \lambda_M[t] \end{bmatrix} \tag{6}$$

where $\boldsymbol{\lambda}[t] = \begin{bmatrix}\lambda_1[t] & \cdots & \lambda_M[t]\end{bmatrix}^\top$, with $\lambda_m[t] = r_m[t] \cdot \mathrm{e}^{j\bar{\omega}_m[t]} \in \mathbb{C}$, $r_m[t] \in \mathbb{R}$, and $\bar{\omega}_m[t] \in [-\pi, \pi]$. The state matrix defined in Equation (6) may be associated with the also diagonal state transition matrix:

$$\boldsymbol{\Phi}[t,\tau] = \prod_{t'=\tau}^{t} \begin{bmatrix}\lambda_1[t'] & & \\ & \ddots & \\ & & \lambda_M[t']\end{bmatrix} = \begin{bmatrix}\rho_1[t,\tau] \cdot \mathrm{e}^{j\varphi_1[t,\tau]} & & \\ & \ddots & \\ & & \rho_M[t,\tau] \cdot \mathrm{e}^{j\varphi_M[t,\tau]}\end{bmatrix} \tag{7}$$

where:

$$\rho_m[t,\tau] = \left|\prod_{t'=\tau}^{t} \lambda_m[t']\right| = \prod_{t'=\tau}^{t} r_m[t'] \qquad \varphi_m[t,\tau] = \arg\left\{\prod_{t'=\tau}^{t} \lambda_m[t']\right\} = \sum_{t'=\tau}^{t} \bar{\omega}_m[t'] \tag{8}$$

are the *instantaneous amplitude* and *instantaneous phase* associated with each state $z_m[t]$, $m = 1, \ldots, M$, related to an excitation at time τ.

In the sequel, it shall be shown how the natural response, the impulse response and the forced response of the diagonal state space system in Equation (6) take the form of the superposition of a set of sinusoids with time-dependent amplitudes and frequencies.

3.1. Modal form of the natural response

Introducing the state transition matrix on Equation (7) into Equation (2b), with initial state $\boldsymbol{z}[0]$ and null excitation $u[t] = 0, \forall t$, yields the natural response of the diagonal state space representation, which leads to the result:

$$\begin{aligned} y[t] &= \begin{bmatrix}c_1[t] & \cdots & c_M[t]\end{bmatrix} \cdot \begin{bmatrix}\rho_1[t,0] \cdot \mathrm{e}^{j\varphi_1[t,0]} & & \\ & \ddots & \\ & & \rho_M[t,0] \cdot \mathrm{e}^{j\varphi_M[t,0]}\end{bmatrix} \cdot \begin{bmatrix}z_1[0] \\ \vdots \\ z_M[0]\end{bmatrix} + w[t] \\ &= \sum_{m=1}^{M} c_m[t] \cdot \rho_m[t,0] \cdot \mathrm{e}^{j\varphi_m[t,0]} \cdot z_m[0] + w[t] \end{aligned} \tag{9}$$

or

$$y[t] = \sum_{m=1}^{M} A_m[t] \cdot \mathrm{e}^{j\varphi_m[t]} + w[t] = \sum_{m=1}^{M} y_m[t] + w[t] \tag{10}$$

where $y_m[t] := A_m[t] \cdot \mathrm{e}^{j\varphi_m[t]}$ is the m–th component of the natural response with *instantaneous amplitude* $A_m[t] \triangleq c_m[t] \cdot \rho_m[t,0] \cdot z_m[0]$ and *instantaneous phase* $\varphi_m[t] \triangleq \varphi_m[t,0]$. Thus, each component of the natural response of the time-dependent system resembles an analytic signal[1]. However, to comply with the strict definition of the analytic signal, the following conditions must hold [9], [1, pp. 35–36]: (*i*) the spectrum of the amplitude component $A_m[t]$ lies in the region $|\omega| < \omega_c$; (*ii*) the spectrum of $\cos\varphi_m[t]$ exists only on the interval $|\omega| \geq \omega_c$.

[1] *Analytic signal*: a complex signal whose spectrum only exists in the positive frequency axis [1, pp. 30].

If the previous conditions hold, then the *Instantaneous Frequency* (IF) of the m–th modal component may be uniquely defined as the time derivative of the respective instantaneous phase [9]. In the presently considered discrete-time case, the following first order backward finite difference approximation of the IF is used instead [40]:

$$\omega_m[t] = \frac{\mathrm{d}\varphi_m[t]}{\mathrm{d}t} \approx \frac{\varphi_m[t] - \varphi_m[t-1]}{T_s} = \frac{\bar{\omega}_m}{T_s} \text{ [rad/s]} \tag{11}$$

Naturally, this would be an effective approximation of the IF as long as the higher order derivatives of the actual IF are small enough. Hence, both the IA and IF are limited to be slowly varying functions of time, according to condition (*i*) of the analytic signal, and the condition on the high order derivatives of the IF.

3.2. Forced response as a superposition of modal filters

The forced response of the diagonal time–dependent state space system with the state matrix defined in Equation (6), evaluated at time t, corresponds to the superposition of the output of M time–dependent modal filters, as follows:

$$z_m[t] = \lambda_m[t] \cdot z_m[t-1] + b_m[t] \cdot u[t] \tag{12a}$$

$$y[t] = \sum_{m=1}^{M} c_m[t] \cdot z_m[t] + w[t] \tag{12b}$$

In turn, each modal filter may be associated with a modal impulse response $h_m[t, t-\tau]$, defined as the response of the system to an impulse at time τ, so that:

$$h_m[t, t-\tau] = \lambda_m[t] \cdot h_m[t-1, t-1-\tau] + b_m[t] \cdot \delta[\tau] \tag{13}$$

As a generalization of the Frequency Response Function (FRF) in the stationary case, the *instantaneous FRF* of a time-dependent system $H[t,\omega)$, is defined as the response of the system to a complex exponential excitation normalized by the excitation itself [35]. Based on this definition, the instantaneous FRF may be computed as the Fourier transform of the time-dependent impulse response over the dimension τ [3, 35], which in the case of each modal filter leads to the following expression for the instantaneous FRF:

$$\begin{aligned} H_m[t,\omega) :=& \mathcal{F}\{h[t, t-\tau]\}_{\tau\to\omega} \\ =& \mathcal{F}\{\lambda_m[t] \cdot h_m[t-1, t-1-\tau] + b_m[t] \cdot \delta[\tau]\}_{\tau\to\omega} \\ =& \lambda_m[t] \cdot H_m[t-1,\omega) \cdot \mathrm{e}^{-j\omega T_s} + b_m[t] \end{aligned} \tag{14}$$

If the time-dependent system is slowly varying, so that $H_m[t,\omega) \approx H_m[t-1,\omega)$, then the instantaneous FRF simplifies into the well-known rational form [3]:

$$H_m^{(F)}[t,\omega) = \frac{b_m[t]}{1 - \lambda_m[t] \cdot \mathrm{e}^{-j\omega T_s}} = \frac{b_m[t]}{1 - r_m[t] \cdot \mathrm{e}^{j(\bar{\omega}_m[t] - \omega \cdot T_s)}} \tag{15}$$

which is referred to as the *frozen FRF*. Note that the only requirement for Equation (15) to be valid is that both the instantaneous eigenvalues and the input coupling coefficients satisfy the conditions $\lambda_m[t] \approx \lambda_m[t-1]$ and $b_m[t] \approx b_m[t-1]$.

The frozen FRF of each modal filter $H_m^{(F)}[t,\omega)$ is characterized by an instantaneous resonance at $\omega = \bar{\omega}_m[t]/T_s$, while its respective bandwidth is determined by $r_m[t]$. In addition, if the excitation is characterized by a single sinusoidal component, say $u[t] = \mathrm{e}^{j\theta[t]}$ with $\theta[t] = \sum_{\tau=0}^{t} \gamma[t]$, then the response would be of the form:

$$y_m[t] = H_m^{(F)}[t, \gamma[t]/T_s) \cdot \mathrm{e}^{j\theta[t]} \tag{16}$$

which implies that the response of the modal filter corresponds to an amplitude modulated version of the sinusoidal excitation. On the other hand, if the excitation is broadband and flat (*v.gr.* white Gaussian noise), then, the instantaneous response would be composed by an infinite set of sinusoids, with higher amplitude on those sinusoids with frequencies closer to the instantaneous resonance.

4. Time-dependent second order block diagonal state space models

The diagonal state space system analyzed in the previous section is characterized by complex coefficients and modal components, which would complicate the practical implementation of modal decomposition algorithms. In this section are discussed alternative modal representations based on block diagonal time-dependent state space representations built from second order blocks, with the advantage that all the coefficients and modal components are real numbers. Initially, second order forms are discussed for the representation of a single modal component, and then multiple second order blocks are used to represent multiple modal components.

4.1. Second order time-dependent state space models

Consider the following diagonal-form state space model:

$$\tilde{\boldsymbol{z}}[t] = \boldsymbol{\Lambda}_m[t] \cdot \tilde{\boldsymbol{z}}[t-1] + \bar{\boldsymbol{b}}_m \cdot u[t] \tag{17a}$$

$$y[t] = \bar{\boldsymbol{c}}_m \cdot \tilde{\boldsymbol{z}}[t] + w[t] \tag{17b}$$

where:

$$\boldsymbol{\Lambda}_m[t] = \begin{bmatrix} \lambda_m[t] & \\ & \lambda_m^*[t] \end{bmatrix} \qquad \bar{\boldsymbol{b}}_m = \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix} \qquad \bar{\boldsymbol{c}}_m = \begin{bmatrix} 1 & 1 \end{bmatrix}$$

where $\lambda_m[t] = r_m[t] \cdot \mathrm{e}^{j\omega_m[t]}$, and the superindex $(^*)$ indicates the complex conjugate.

According to Equation (10), the natural response of the state space representation in Equation (17) is of the form:

$$\begin{aligned} y[t] &= \frac{1}{2}\left(\rho_m[t] \cdot \mathrm{e}^{j\varphi_m[t]} + \rho_m[t] \cdot \mathrm{e}^{-j\varphi_m[t]}\right) + w[t] \\ &= \frac{1}{2} \cdot \rho_m[t] \cdot \cos\varphi_m[t] + w[t] \end{aligned} \tag{18}$$

where $\rho_m[t] = \prod_{\tau=1}^{t} r_m[\tau]$ and $\varphi_m[t] = \sum_{\tau=1}^{t} \bar{\omega}_m[\tau]$.

The state space representation in Equation (17) may be transformed into an equivalent state space representation where all the variables are real, through the *time–invariant* transformation matrix:

$$\boldsymbol{V} = \begin{bmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ j & -j \end{bmatrix}, \qquad \boldsymbol{V}^{-1} = \boldsymbol{U} = \begin{bmatrix} \boldsymbol{u}_1 & \boldsymbol{u}_2 \end{bmatrix} = \frac{1}{2} \cdot \begin{bmatrix} 1 & -j \\ 1 & j \end{bmatrix} \tag{19}$$

where $\boldsymbol{v}_k$ are the time–invariant right eigenvectors, and $\boldsymbol{u}_k$ are the time–invariant left eigenvectors of the matrix $\boldsymbol{\Lambda}_m[t]$ at time t. Applying these transformation matrices into the representation in Equation (17), leads to the symmetric state space representation:

$$\boldsymbol{z}[t] = \boldsymbol{M}_m[t] \cdot \boldsymbol{z}[t-1] + \boldsymbol{b}_m \cdot u[t] \tag{20a}$$

$$y[t] = \boldsymbol{c}_m \cdot \boldsymbol{z}[t] + w[t] \tag{20b}$$

where:

$$\boldsymbol{M}_m[t] = \begin{bmatrix} \alpha_m[t] & \beta_m[t] \\ -\beta_m[t] & \alpha_m[t] \end{bmatrix}, \qquad \boldsymbol{b}_m = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad \boldsymbol{c}_m = \begin{bmatrix} 1 & 0 \end{bmatrix} \tag{21}$$

characterized by the eigenvalues $\lambda_m[t] = \alpha_m[t] + j\beta_m[t] = r_m[t] \cdot \mathrm{e}^{j\bar{\omega}_m[t]}$, where $\alpha_m[t]$ and $\beta_m[t]$ correspond to the real and imaginary parts of the instantaneous eigenvalues. If the term $r_m[t] \equiv 1$ for all t, then the response of the system will have only sinusoidal components of the form $\cos\bar{\varphi}_m[t,\tau]$. Moreover, $\alpha_m[t] = \cos\bar{\omega}_m[t]$ and $\beta_m[t] = \sin\bar{\omega}_m[t]$, rendering the state matrix into the form:

$$\boldsymbol{M}_m[t] = \begin{bmatrix} \cos\bar{\omega}_m[t] & \sin\bar{\omega}_m[t] \\ -\sin\bar{\omega}_m[t] & \cos\bar{\omega}_m[t] \end{bmatrix} \tag{22}$$

The time–dependent symmetric state space representation of Equation (21) is endowed with some appealing properties, which are listed next:

(i) The transformation matrix $\boldsymbol{V}$ and its respective eigenvectors are time–invariant. This means that the transformation between the diagonal and symmetric representations is always well–defined, since the transformation matrix and its inverse are also well–defined.

(ii) The state vector contains the in–phase modal component and its quadrature, thus directly providing an analytic representation of the modal component. To show this property, consider first the state transition matrix of the symmetric representation, which, according to Equation (5), is as follows:

$$\begin{aligned} \boldsymbol{\Phi}[t,\tau] &= \boldsymbol{V} \cdot \bar{\boldsymbol{\Phi}}[t,\tau] \cdot \boldsymbol{V}^{-1} \\ &= \frac{1}{2} \cdot \rho_m[t,\tau] \cdot \begin{bmatrix} 1 & 1 \\ j & -j \end{bmatrix} \cdot \begin{bmatrix} \mathrm{e}^{j\varphi_m[t,\tau]} & \\ & \mathrm{e}^{j\varphi_m[t,\tau]} \end{bmatrix} \cdot \begin{bmatrix} 1 & -j \\ 1 & j \end{bmatrix} \\ &= \rho_m[t,\tau] \cdot \begin{bmatrix} \cos\varphi_m[t,\tau] & \sin\varphi_m[t,\tau] \\ -\sin\varphi_m[t,\tau] & \cos\varphi_m[t,\tau] \end{bmatrix} \end{aligned} \tag{23}$$

Then, the natural solution of the state vector for initial value $\boldsymbol{z}[0] = \begin{bmatrix} z[0] & 0 \end{bmatrix}^T$ at time t is given by Equation (2), which in the present case becomes:

$$\begin{aligned} \boldsymbol{z}[t] &= \boldsymbol{\Phi}[t,0] \cdot \boldsymbol{z}[0] = \rho_m[t,0] \cdot \begin{bmatrix} \cos\varphi_m[t,0] & \sin\varphi_m[t,0] \\ -\sin\varphi_m[t,0] & \cos\varphi_m[t,0] \end{bmatrix} \cdot \begin{bmatrix} z[0] \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} \rho_m[t,0] \cdot z[0] \cdot \cos\varphi_m[t] \\ -\rho_m[t,0] \cdot z[0] \cdot \sin\varphi_m[t] \end{bmatrix} = \begin{bmatrix} z_1[t] \\ z_2[t] \end{bmatrix} \end{aligned} \tag{24}$$

Thus, it is clear that the state $z_1[t]$ is related to cosine component, while the state $z_2[t]$ is related to sine component, both with the same frequency and amplitude.

(iii) Since both states constitute the analytic form of the response, then the instantaneous amplitude and phase are determined from the state vector through the relations:

$$A_m[t] = (z_1^2[t] + z_2^2[t])^{1/2}; \qquad \varphi_m[t] = \arctan\left(\frac{z_2[t]}{z_1[t]}\right) \tag{25}$$

(iv) The representation is fully characterized by the time–dependent parameter vector $\boldsymbol{\theta}[t] = \begin{bmatrix} \alpha_m[t] & \beta_m[t] \end{bmatrix}^\top$.

4.2. M-order block-diagonal time-dependent state space representations

A signal conformed by M sinusoidal components can be represented by the following block–diagonal representation:

$$\boldsymbol{M}[t] = \begin{bmatrix} \boldsymbol{M}_1[t] & & & \\ & \boldsymbol{M}_2[t] & & \\ & & \ddots & \\ & & & \boldsymbol{M}_M[t] \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} \boldsymbol{b}_1 \\ \boldsymbol{b}_2 \\ \vdots \\ \boldsymbol{b}_M \end{bmatrix}, \quad \boldsymbol{c} = \begin{bmatrix} \boldsymbol{c}_1 & \boldsymbol{c}_2 & \cdots & \boldsymbol{c}_M \end{bmatrix} \tag{26}$$

where each of the blocks $\boldsymbol{M}_m[t]$, $\boldsymbol{b}_m$ and $\boldsymbol{c}_m$ are defined in Equation (21). This block-diagonal state space representation may be associated with the natural response:

$$y[t] = \sum_{m=1}^{M} A_m[t] \cdot \cos\varphi_m[t] + w[t] \tag{27}$$

while the instantaneous amplitude and phase may be derived from Equation (25) on each one of the modal components. The obtained block diagonal symmetric state space representation is characterized by the number of modal components M, the innovations variance $\sigma_w^2[t]$ and the parameter vector $\boldsymbol{\theta}[t] \in \mathbb{R}^{2M}$, the latter defined as:

$$\boldsymbol{\theta}[t] = \begin{bmatrix} \alpha_1[t] & \beta_1[t] & \alpha_2[t] & \beta_2[t] & \cdots & \alpha_M[t] & \beta_M[t] \end{bmatrix}^\top \tag{28}$$

5. Kalman filter estimation methods of time-dependent modal decompositions

5.1. Concurrent estimation of modal trajectories and modal parameters

As explained in Section 3, the natural response of a diagonal state space system takes the form of a superposition of modal components with time-dependent amplitude and frequency. Furthermore, as shown in Section 4, particular block diagonal structures may also be directly associated with the response of a diagonal state space system through a time-invariant equivalence transformation. In this section, it shall be shown how block diagonal structures can be used to estimate a modal decomposition of a given non-stationary signal.

Hence, consider the non-stationary signal $y[t]$, which is assumed to comply with an M-order modal decomposition of the form of Equation (10). In order to estimate the modal trajectories, IA and IF, the following block diagonal state space system is appraised:

$$\boldsymbol{z}[t] = \boldsymbol{M}(\boldsymbol{\theta}[t]) \cdot \boldsymbol{z}[t-1] + \boldsymbol{b} \cdot u[t], \qquad u[t] \sim \texttt{NID}\left(0, \sigma_u^2\right) \tag{29a}$$

$$y[t] = \boldsymbol{c} \cdot \boldsymbol{z}[t] + w[t], \qquad w[t] \sim \texttt{NID}\left(0, \sigma_w^2\right) \tag{29b}$$

where $\boldsymbol{M}(\boldsymbol{\theta}[t]) \equiv \boldsymbol{M}[t]$, $\boldsymbol{b}$ and $\boldsymbol{c}$ are defined in Equation (26), and $u[t]$ and $w[t]$ are zero mean NID noise processes, with respective variances σ_u^2 and σ_v^2. It must be remarked that, although the aim is to associate the observed signal with the natural response of the state space system, the NID noise excitation $u[t]$ is preserved within the estimation model to account for the uncertainty in the modal trajectories.

One more equation must be introduced into the state space system to describe the evolution of the parameter vector $\boldsymbol{\theta}[t]$. Presently, considering the slow variation assumption of the dynamics, the following stochastic first difference model is appraised:

$$\boldsymbol{\theta}[t] = \boldsymbol{\theta}[t-1] + \boldsymbol{v}[t], \qquad \boldsymbol{v}[t] \sim \texttt{NID}\left(\boldsymbol{0}_d, \boldsymbol{\Sigma}_v\right) \tag{30}$$

where d is the dimension of the parameter vector, and $\boldsymbol{v}[t]$ is the NID parameter innovations process with mean zero and covariance matrix $\boldsymbol{\Sigma}_v \in \mathbb{R}^{d \times d}$. Here, it is assumed that the parameter innovations covariance matrix is of the form $\boldsymbol{\Sigma}_v = \sigma_v^2 \cdot \boldsymbol{I}$, so as to reduce the number of fixing parameters of the method.

Given the state space representation of Equation (29) and the parameter evolution model in Equation (30), the problem of estimation of the modal trajectories and modal parameters translates into the concurrent estimation of both state and parameters, which may be solved through a *Joint Kalman Filter* (JKF) or *Dual Kalman Filter* (DKF) approach. In the *Joint Kalman Filter Non-stationary Sinusoid Tracking* (JKF-NST) method, both state and parameters are collectively represented in terms of a joint state vector defined as $\boldsymbol{z}_J[t] = \left[\boldsymbol{z}^T[t] \quad \boldsymbol{\theta}^T[t]\right]^T$, and then a single Kalman filter is used for state estimation. On the other hand, the *Dual Kalman Filter Non-stationary Sinusoid Tracking* (DKF-NST) method is based directly on the concurrent state space representation on equations (29) and (30), while estimation is performed by two coupled Kalman filters, each one estimating the state and the parameter vector. Details of these estimation methods can be found in [41].

A simplified version of the JKF-NST method is achieved by utilizing the simplified blocks described in Equation (22), thus leading to the methods known as *Kalman Filter*

Frequency Trackers (KF-FT), where estimation is hinged either on *Extended Kalman filters* [27, 28], or on more robust non-linear state estimation methods, such as *Cubature or Unscented Kalman filters* [31, 32, 30]. In the KF-FT methods, the number of parameters to estimate is reduced, but at the same time, a stronger type of non-linearity is introduced in the state space equations, which can make estimation more sensitive to noise or perturbations.

5.2. TAR-model based approach

A TAR(n_a) model, with n_a designating its AR order, is defined as follows [3]:

$$y[t] = -\sum_{i=1}^{n_a} a_i[t] \cdot y[t-i] + u[t], \qquad u[t] \sim \texttt{NID}\left(0, \sigma_u^2[t]\right) \tag{31}$$

where $a_i[t]$ are the *time-dependent AR parameters*, and $u[t]$ an (unobservable) normally and identically distributed *innovations process* with mean zero and variance $\sigma_u^2[t]$. The TAR model in Equation (31) may be cast into the time-dependent state space representation shown in Equation (1), where:

$$\boldsymbol{A}_{TAR}[t] = \begin{bmatrix} -a_1[t] & -a_2[t] & \cdots & -a_{n_a-1}[t] & -a_{n_a}[t] \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}_{n_a \times n_a} \qquad \boldsymbol{b} = \boldsymbol{c}^T = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{32}$$

Then, a modal decomposition may be obtained from the state space form of the TAR model by means of a proper equivalence transform, as discussed in Section 2. A necessary (but not sufficient) condition to obtain a representation consisting of M sinusoidal components, as in Equation (26), is to use a TAR model of order $n_a = 2 \cdot M$. The transformation matrix $\boldsymbol{V}[t]$ is calculated from the following instantaneous eigenvalue decomposition:

$$\boldsymbol{A}_{TAR}[t] \cdot \lambda[t] = \lambda[t] \cdot \boldsymbol{v}[t] \tag{33}$$

Hence, under the assumption that the parameters of the TAR model are slowly varying and that no eigenvalues are repeated, the original state space representation is transformed into a diagonal one, through the approximate equivalence transformation:

$$\boldsymbol{\Lambda}[t] = \boldsymbol{V}[t] \cdot \boldsymbol{A}_{TAR}[t] \cdot \boldsymbol{V}^{-1}[t-1] \approx \boldsymbol{V}[t] \cdot \boldsymbol{A}_{TAR}[t] \cdot \boldsymbol{V}^{-1}[t] \tag{34}$$

and similarly for the matrices $\boldsymbol{b}$ and $\boldsymbol{c}$.

This type of modal decomposition has been analyzed in [33, 34, 42], although its application meets some practical difficulties. One of these is that the number of instantaneous real and complex conjugate roots cannot be controlled, and as a result, these quantities may shift between time instants. Consequently, discontinuities in the extracted modes are typically found. In addition, the values of the roots are not tracked directly and for that reason, these should be sorted and matched with their values on previous time instants. As a result, the estimated values of the modal trajectories often exhibit jumps or discontinuities.

5.3. Initialization of the modal decomposition methods

Initial values of the state and parameter vector estimates ($\hat{\boldsymbol{z}}[1]$ and $\hat{\boldsymbol{\theta}}[1]$) can be derived from an AutoRegressive (AR) model of order $n_a = 2 \cdot M$ estimated from on a short initial fragment of the signal. Then the initial values of the modal trajectories may be obtained as:

$$\hat{\boldsymbol{z}}[1] = \boldsymbol{V} \cdot \boldsymbol{W}^{-1} \cdot \boldsymbol{y}[1], \quad \boldsymbol{y}[1] = \begin{bmatrix} y[1] & 0 & \cdots & 0 \end{bmatrix}^{\top}_{1 \times 2M} \tag{35}$$

where $\boldsymbol{A}_{AR} = \boldsymbol{W} \cdot \boldsymbol{\Lambda} \cdot \boldsymbol{W}^{-1}$ is the eigenvalue decomposition of the state matrix associated with the short fragment estimated AR model, and

$$\boldsymbol{V} = \boldsymbol{I}_M \otimes \begin{bmatrix} 1 & 1 \\ j & -j \end{bmatrix} \tag{36}$$

is a transformation matrix that maps the complex diagonal state space representation into a real block-diagonal one. In addition, $\hat{\boldsymbol{\theta}}[1]$ can be initialized from the entries in the diagonal eigenvalue matrix $\boldsymbol{\Lambda}$.

5.4. Optimization of the adjustment parameters

The adjustment parameters of the modal decomposition methods are the variances σ_u^2, σ_v^2 and σ_w^2. These parameters control the tracking ability of the Kalman filter and noisiness of the estimates. More precisely, when the ratio σ_v^2/σ_w^2 is large then the Kalman filter adjusts faster to the changes in the parameters but may yield noisier estimates. Likewise, when the ratio σ_u^2/σ_w^2 is large, then the Kalman filter adjusts faster to the changes in the modal components. Otherwise, small values of these ratios would lead to less variability in the estimates, but may hinder the tracking ability.

Although heuristic methods may be used to adjust these variances, optimization based on the likelihood or prediction error is more recommendable. The *Marginal Likelihood method* associated with the Kalman filter is defined as [43]; [44, Ch. 6]; [45, Ch. 6]:

$$\ln \mathcal{L}(\sigma_u^2, \sigma_v^2, \sigma_w^2) = K - \frac{1}{2} \sum_{t=2}^{N} \left(\ln \hat{\sigma}_e^2[t] + \frac{\hat{w}^2[t|t-1]}{\hat{\sigma}_e^2[t]} \right) \tag{37a}$$

$$\hat{w}[t|t-1] = y[t] - \boldsymbol{c} \cdot \hat{\boldsymbol{z}}[t|t-1] \tag{37b}$$

$$\hat{\sigma}_e^2[t] = \boldsymbol{c} \cdot \boldsymbol{P}[t|t-1] \cdot \boldsymbol{c}^{\top} + \sigma_w^2 \tag{37c}$$

where $\hat{\boldsymbol{z}}[t|t-1]$ is the *a priori* estimate of the state vector and:

$$\boldsymbol{P}[t|t-1] = \mathrm{E}\left\{ (\boldsymbol{z}[t] - \hat{\boldsymbol{z}}[t|t-1]) \cdot (\boldsymbol{z}[t] - \hat{\boldsymbol{z}}[t|t-1])^{\top} \right\} \tag{38}$$

is the error covariance matrix of the *a priori* state estimate, both provided by the KF.

On the other hand, the *Mean Squared Prediction Error* (MSPE), defined as:

$$MSPE(\sigma_u^2, \sigma_v^2, \sigma_w^2) = \frac{1}{N} \cdot \sum_{t=1}^{N} \hat{w}^2[t|t-1] \tag{39}$$

where $\hat{w}[t|t-1]$ denotes the prediction error, provides a simpler method to compute the performance of the modal decomposition.

Optimization based on any of the above mentioned objective functions may be tricky due to the presence of local minima. In addition, in some cases numerical instabilities appear for some values of the variances. To cope with this problem it is recommendable to make a coarse search in the parameter space before attempting to use a numerical optimization scheme. Moreover, to further simplify the optimization problem, the number of optimized parameters may be reduced by defining $\sigma_w^2 = 1$, and then optimizing only σ_u^2 and σ_v^2.

6. Benchmark test

The following test aims at comparing the performance of the proposed modal decomposition methods in the decomposition of a non–stationary signal composed by three time-dependent sinusoidal components embedded in white Gaussian noise. The time-frequency characteristics of the signal include mode crossing and temporary fading of one of the modal components. The four methods discussed in Section 5 are compared, namely the JKF-NST, DKF-NST, EKF-FT, and the TAR model–based Modal Decomposition (TAR-MD) methods. In addition, two well–known TFR-based methods based on ridge extraction are considered for comparison [6, 8]: the first one based on the instantaneous maximum of the TFR (TFR-maximum); the second one based on a ridge extraction scheme with adaptive penalization of frequency jumps (TFR-adaptive or *scheme II*(α,β) as described in [8]). Details are shown next.

6.1. The signals

The non-stationary signals used in the test are composed by three time-dependent sinusoidal components, whose IA and IF laws are defined as follows:

$$A_1[t] = A_{1,0} + A_{1,1} \cdot \cos 2\pi T_s \alpha_1 t \qquad f_1[t] = f_{1,0} + f_{1,1} \cdot \sin 2\pi T_s \alpha_1 t \tag{40a}$$

$$A_2[t] = A_{2,0} + A_{2,1} \cdot \sin 2\pi T_s \alpha_2 t \qquad f_2[t] = f_{2,0} + f_{2,1} \cdot \sin 2\pi T_s \alpha_2 t \tag{40b}$$

$$A_3[t] = \max(A_{3,0} + A_{3,1} \cdot |\cos 2\pi T_s \alpha_3 t|, 0) \qquad f_3[t] = f_{3,0} + f_{3,1} \cdot \sin 2\pi T_s \alpha_1 t \tag{40c}$$

where $t = 1, 2, \ldots, N$ is the normalized discrete time and T_s is the sampling period. The coefficients $A_{m,0}$ and $f_{m,0}$ for $m = 1,2,3$ indicate the mean amplitude and frequency, while the coefficients $A_{m,1}$ and $f_{m,1}$ for $m = 1,2,3$ indicate the variation of the amplitude and frequency, and the coefficients α_m for $m = 1,2,3$ indicate the normalized frequencies of the amplitude and frequency variations. The values of these coefficients used in the simulations are summarized in Table 1. Each non-stationary signal is then constructed as follows:

$$y[t] = \sum_{m=1}^{M} A_m[t] \cdot \mathrm{e}^{j\varphi_m[t]} + w[t], \qquad w[t] \sim \texttt{NID}\left(0, \sigma_w^2\right) \tag{41a}$$

$$\varphi_m[t] = \varphi_{m,o} + 2\pi T_s \sum_{\tau=1}^{t} f_m[t] \tag{41b}$$

where $\varphi_{m,o} \in [0, 2\pi]$ is the random uniformly distributed initial phase, and $w[t]$ is a zero mean NID noise with variance σ_w^2. On a single simulation, a signal of length $N = 3\,000$ samples is created according to the previously described IA and IF laws, based on an independent realization of the initial phase $\varphi_{m,0}$ and the noise $w[t]$, while the noise variance σ_w^2 is adjusted according to the desired SNR value. Figure 1 shows a single realization of the signal. The signal is characterized by the following features: crossing of modes 1 and 2 at time $t = 1.5$ *s*; near crossing of modes 2 and 3 around time $t = 2.5$ *s*; and vanishing amplitude of mode 3 at $t = 1$ *s*.

Table 1: Signal and simulation parameters applied in the benchmark test

Mode	$A_{m,0}$	$A_{m,1}$	$f_{m,0}$	$f_{m,1}$	α_m
$m = 1$	1.0	0.2	50	-20	0.50
$m = 2$	0.8	0.4	80	20	0.50
$m = 3$	1.0	-0.1	120	20	0.25
Sampling period	$T_s = 10^{-3}$ s				
Signal length	$N = 3\,000$ samples				
Number of tests	$N_{MC} = 40$ tests				
SNR values	$SNR = \{-20, -15, -10, \ldots, 20\}$ dB				

6.2. Description of the methods

All the methods are applied assuming that the number of modal components is known. A description of each method is provided next:

JKF-NST, DKF-NST and EKF-FT methods: Initial values are extracted from an AR(6) model estimated from the initial 200 samples of the signal. Optimal performance is pursued by a *Genetic Algorithm* (GA) operating on the MSPE, defined in Equation (39). The optimized variables are the state and parameter variances σ_u^2 and σ_v^2, and the initial error variance p_0, so that the initial error covariance matrix is of the form $\boldsymbol{P}[0|0] = p_0 \cdot \boldsymbol{I}$. Moreover, the measurement error variance σ_w^2 is defined equal to one, so as to reduce the number of optimized parameters. The variances are optimized for each SNR value based on a single realization of the signal.

TAR-MD method: This method is adjusted according to the guidelines provided in [3], using a *second order smoothness priors* TAR modeling approach. The optimal order of the TAR model is selected by minimizing the Bayesian Information Criterion, which in the present case leads to the model order $n_a = 18$. Several spurious modes appear in the TAR model based modal decomposition, since the model order is much larger than twice the number of components. Therefore, only the first three modes with lowest damping are considered as the actual modes of the signal, while the remaining ones are considered as spurious modes.

TFR-maximum: Achieved by locating the maximum of the TFR. Multiple components are extracted by the *peeling* method. TFR estimates obtained with the spectrogram (windowed Fourier Transform) based on a Gaussian window with resolution parameter $f_0 = 0.05$ Hz. Implementation is achieved through the MATLAB toolbox suggested in [8].

TFR-adaptive: Corresponds to the curve extraction method with adaptive penalization of frequency jumps suggested in [8], where is referred to as *scheme II*(α, β).

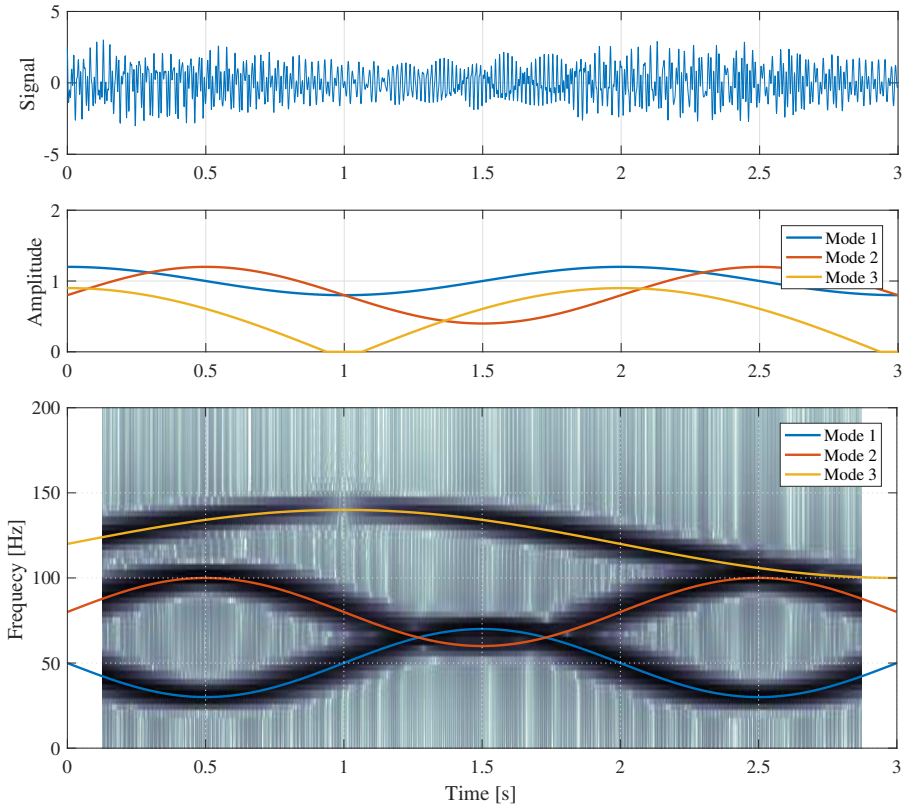


Figure 1: Typical realization of the simulated noise-free non-stationary signal. Top: signal in time domain; Middle: IA of the modal components; Bottom: spectrogram of the signal with overlaid IF of each modal component. Spectrogram parameters: Hamming window, length 256 samples, 255 sample overlap.

The adjustment parameters of the method, namely α and β are both set to 1 as suggested in [8]. TFR estimates obtained with the spectrogram (windowed Fourier Transform) based on a Gaussian window with resolution parameter $f_0 = 0.05$ Hz. Implementation is achieved through the MATLAB toolbox suggested in [8].

6.3. Monte Carlo analysis of the modal decomposition methods

A Monte Carlo test is performed to evaluate the performance of the modal decomposition methods in the estimation of the modal components of the non-stationary signal and their respective IA and IF. SNR values are considered in the range from -20 to 20 dB, while 40 independent Monte Carlo runs are performed at each SNR level. Figure 2 and Figure 3 show the median and the sample distribution of the IF and IA estimates obtained with each one of the modal decomposition methods after the Monte Carlo analysis for an SNR=20 dB. In addition, Figure 4 shows the median MSE value calculated for each one of the modal component estimates and respective IA and IF estimates for the considered SNR range obtained from the 40 Monte Carlo runs. In

general it is observed that the JKF-NST method consistently provides the best performance, although other methods may provide better performance for particular SNR ranges. In addition, it is clear that the level of noise has a detrimental effect in all of the modal decomposition methods. Particularly, for components 1 and 2 characterized by a frequency crossing, the obtained estimates have larger errors. A concise analysis of the performance of each one of the modal decomposition methods is given next:

JKF-NST method: Yields consistently good performance figures among the analyzed methods. As observed in Figure 2(a), it appears to be the least affected by the mode crossing, although it still bears an important influence in the modal amplitude estimates. On the other hand, the vanishing modal component appears to be correctly tracked by the algorithm. In normal operation, the JKF-NST method tracks with very low error each one of the modal components.

DKF-NST method: The overall performance of this method is roughly similar to that of the JKF-NST method in the amplitude and modal trajectory estimates, although the results on the IF are not as robust. As can be observed in Figure 2(b), the reduced performance may be originated by an apparent delay in the IF tracking and the prolonged disturbances originated after the frequency crossing.

EKF-FT method: Similar to the JKF-NST method, the EKF-FT method demonstrates very good tracking of the modal components and their properties, until the frequency crossing event. After that, in many cases the IF estimates loose track of the actual IF, and as an effect, the obtained estimates also bear large differences with the actual values. As a result, the overall performance outcomes displayed in Figure 4 are reduced in contrast to that obtained with the JKF-NST method. Remarkably, it appears to provide very good performance under very low SNR values.

TAR-MD method: As seen in Figure 3(a), the TAR-MD method provides very reliable IF estimates during the frequency crossing event, although in general the obtained IF estimates appear to be biased. However, the IA and modal trajectory estimates have large errors, which make this the least accurate method among the state space modal decomposition methods.

TFR-maximum and TFR-adaptive methods: According to Figure 4, both methods are very reliable in the estimation of the modal components and their amplitude in low SNRs, being of course the TFR-adaptive method the most effective of all. However, the IF estimates are not as precise as those obtained with the JKF-NST method. As can be seen in Figure 3(b)-(c), the mode crossing has a huge detrimental effect in the IF estimates even for the curve extraction method with adaptive penalization (TFR-adaptive), which explains the poor IF estimation performance curves. Otherwise, the methods appear to yield proper estimates.

7. Application example: powerline interference denoising on ECG signals

7.1. The QT ECG database

This application example consists on denoising (removal) of the powerline interference on *ElectroCardio Graphic* (ECG) signals, as a mean to demonstrate the applicability of the modal decomposition methods proposed in this work on signals characterized by complex non-stationary behavior. To that end, ECG records from the publicly

available QT database are extracted and then artificially contaminated with powerline interference. The QT database includes a total of 105 ECG records selected from a large class of QRS and ST–T segment morphologies, made to test QT detection algorithms under similar real–life conditions [46]. Each recording is sampled at 250 Hz and has a length of 15 minutes, and is characterized by sudden changes in amplitude and frequency. Figure 5 shows a segment of typical noise-free signal from the QT database ("sel117"), where the complex characteristics of the signal, including fast frequency and amplitude changes, are evident.

7.2. Modal decomposition of a single noise-free ECG signal

The JKF-NST, DKF-NST and EKF-FT modal decomposition methods are utilized to obtain a modal decomposition of the ECG signal "sel117" shown in Figure 5. Again, the TFR-adaptive method is used to contrast the modal decomposition results. The TAR-based modal decomposition method is not included in this demonstration, given the poor results obtained in the benchmark test, particularly regarding to the estimation capabilities of the modal components.

A total number of 30 modal components is selected for the KF-based modal decomposition algorithms, which corresponds the minimum of the MSPE (Equation (39)) obtained in the range from 1 to 50. The variances are selected heuristically as follows: $\sigma_w^2 = 1$, $\sigma_u^2 = 10^{-3}$, $\sigma_v^2 = 10^{-7}$ and $p_0 = 10^{-7}$. In addition, the parameter vector $\boldsymbol{\theta}[1]$ is initialized with values close to the heart rate observed in the signal and its harmonics. In the case of the TFR-adaptive method, the fixing parameters of the method are selected as in Section 6, differing only on the value of the resolution parameter of the Gaussian window, which in the present case is set to $f_0 = 1$ Hz.

The modal decomposition results are shown in Figures 6 and 7. Left plots show the first 6 modal components, while the right plots show the corresponding IF of each one of the non-stationary modal components overlaid over the spectrogram. The results demonstrate that the KF-based modal decomposition methods provide similar results, which are similarly affected by the sudden frequency change after time $t = 110$ s. The first identified modal component corresponds to the baseline wander, namely the trend observed on the ECG signal. The next component may be associated to the heart rate, or in other words, the fundamental period of the ECG signal. Other components corresponds to harmonics of the heart rate. It is evident that the IF estimates are influenced by the transients corresponding to the QRS segment (largest peak) of the ECG signal.

The TFR-adaptive method identifies almost the same modes as in the case of the KF-based modal decomposition methods, although it is unable to identify the baseline wander component, and instead introduces a spurious mode at about 2 Hz. However, this method is not affected by the sudden frequency change occurring around 110 s. On the other hand, larger oscillations in the IF estimates are observed in contrast to the JKF-NST method.

7.3. Powerline filtering

For the denoising test, each of the original ECG signals from the QT database is artificially contaminated with a non-stationary powerline interference, according to:

$$y_{noisy}[t] = y_{ecg}[t] + y_{pl}[t] \tag{42}$$

$$y_{pl}[t] = \gamma \cdot \sin\left(\frac{2\pi f_{pl}}{f_s} t\right) \cdot \frac{1}{1+\exp(-0.5(t-60)/f_s)} \tag{43}$$

where $y_{ecg}[t]$ is the clean ECG signal, $y_{noisy}[t]$ the noisy ECG signal, and $y_{pl}[t]$ the non-stationary power line interference, with frequency $f_{pl} = 50$ Hz and time-dependent amplitude determined by the coefficient γ. This definition for the powerline interference is used to simulate the onset of the interference at time $t = 60$ seconds. The coefficient γ is used to adjust the SNR of the noisy ECG.

Filtering of the powerline interference is made according to the following procedure: (i) a 30-th order modal decomposition is calculated with the JKF-NST, DKF-NST and EKF-FT methods; (ii) the estimated modal components are analyzed to detect the presence of the powerline interference (50 Hz component), (iii) a denoised ECG signal is obtained by removing the estimated powerline noise component from the noisy ECG signal. As in the modal decomposition performed in the noise-free signal presented in the previous section, the parameter vector $\boldsymbol{\theta}[1]$ of the modal decomposition methods is initialized with values near to the heart rate and its harmonics. In addition, one of the modes is assigned to the frequency of the powerline interference.

A Monte Carlo test is considered to evaluate the performance of the methods, and consists on artificially adding powerline interference to the complete set of ECG records of the QT database for different values of SNR (SNR input), using the interference model of Equation (42). The powerline interference filtering approach based on the proposed state space modal decomposition methods is compared with a denoising method based on the discrete wavelet transform, which is considered as reference in the state-of-the-art [47, 48]. The wavelet denoising method is based on a 10 level discrete Meyer wavelet decomposition. A hard threshold is applied on the higher detail levels (9 and 10) and subsequently, the signal is reconstructed.

The performance of the methods is measured in terms of the MSE and SNR of the ECG signals after having removed the powerline interference component (SNR output). Figure 8 shows the results of this test using input SNR values in the range $SNR_{in} = [-40, \ldots, 40]$ dB. The outcomes demonstrate that in general the SNR curves are characterized by a linear increment until a saturation point is achieved. Such saturation point is much higher for the KF-based methods and is consistently found near to 32 dB. In contrast, the DWT-based method saturates at about 22 dB. This indicates that the KF-based methods provide better performance in comparison to the DWT method. In addition, the SNR output values are higher in the case of the JKF-NST and DKF-NST methods, which indicates an improved reduction of the powerline interference even under low SNR values. From these methods, the JKF-NST-based method is the one consistently yielding the highest SNR gain. The saturation of the SNR gain is an indication of the information removed from the original signal by the denoising method. This means, in the case of the KF-based methods, that the removed powerline modal component also contains information of the ECG signal. However, the amount

of information removed by the KF-based methods is reduced in contrast to the DWT method.

8. Conclusions

The problem of modal decomposition of non-stationary signals based on diagonal linear time-dependent state space representations has been discussed in this work. As an initial contribution, it is demonstrated with the use of equivalence transforms that the response of a linear time-dependent state space representation can be described as the superposition of a finite number of narrowband frequency components, which in the case of the natural response reduce to modal components with time-dependent amplitude and frequency. In addition, based on the previous result, a particular block diagonal state space structure is proposed and applied to compute modal decompositions of non-stationary signals in combination with Kalman filters. Two Kalman filter based methods are proposed to estimate modal decompositions, one hinged on joint estimation of states and parameters of the block diagonal representation (JKF-NST), the other hinged on decoupled estimation of states and parameters (DKF-NST). Both methods correspond to generalizations of the well-known Kalman Filter Frequency Trackers (KF-FT), with the main advantage being that the obtained methods are linear-in-the-parameters. Analysis carried out in a Monte Carlo test on a signal featuring mode crossing and mode vanishing, and in the modal decomposition and denoising of ECG records demonstrate the increased tracking accuracy and reliability of the proposed methods, particularly in the case of the JKF-NST method.

As a final conclusion, it can be stated that the proposed methods constitute a reliable tool for on-line modal decomposition of multi-component non-stationary signals, with accurate results comparable and even better than other state-of-the-art methods. Further efforts can be directed towards the refinement of the proposed methods for special cases, including for example the case of harmonics or the case of vanishing/appearing modal components. Likewise, it would be of interest to research methods to self-tune the modal decomposition methods as they operate on a non-stationary signal, thus avoiding the issue of adjusting of the process and measurement noise variances.

Acknowledgements

Luis David Avendaño-Valencia would like to acknowledge the support of the ETH Zurich Postdoctoral Fellowship FEL-45 14-2 *"A data-driven computational framework for damage identification and life-cycle management of wind turbine facilities"*.

References

[1] L. Cohen, Time-Frequency Analysis, Prentice Hall Signal Processing Series, Prentice Hall, 1995.

[2] S. Mallat, A Wavelet Tour of Signal Processing, 3rd Edition, Academic Press, 2008.

[3] A. Poulimenos, S. Fassois, Parametric time-domain methods for non-stationary random vibration modeling and analysis: A critical survey and comparison, Mechanical Systems and Signal Processing 20 (4) (2006) 763–816.

[4] B. Boashash (Ed.), Time Frequency Signal Analysis and Processing, Elsevier, 2003.

[5] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, H. H. Liu, The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis, Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences 454 (1971) (1998) 903–995.

[6] I. Djurović, Viterbi algorithm for chirp-rate and instantaneous frequency estimation, Signal Processing 91 (2011) 1308–1314.

[7] L. Stanković, I. Djurović, S. Stanković, M. Simeunović, S. Djukanović, M. Daković, Instantaneous frequency in time-frequency analysis: Enhanced concepts and performance of estimation algorithms, Digital Signal Processing 35 (2014) 1–13.

[8] D. Iatsenko, P. V. E. McClintock, A. Stefanovska, Extraction of instantaneous frequencies from ridges in time-frequency representations of signals, Signal Processing 125 (2016) 290–303.

[9] B. Boashash, Estimating and interpreting the instantaneous frequency of a signal – Part 1: Fundamentals, Proceedings of the IEEE 80 (4) (1992) 520–538.

[10] Y. Qin, Y. Mao, B. Tang, Multicomponent decomposition by wavelet modulus maxima and synchronous detection, Mechanical Systems and Signal Processing 91 (2017) 57–80.

[11] L. Rankine, M. Mesbah, B. Boashash, IF estimation for multicomponent signals using image processing techniques in the time-frequency domain, Signal Processing 87 (2007) 1234–1250.

[12] H. Li, Z. Li, W. Mo, A time varying filter approach for empirical mode decomposition, Signal Processing 138 (2017) 146–158.

[13] E. Alickovic, J. Kevric, A. Subasi, Performance evaluation of empirical mode decomposition, discrete wavelet transform and wavelet packed decomposition for automated epileptic seizure detection and prediction, Biomedical Signal Processing and Control 39 (2018) 94–102.

[14] M. Blanco-Velasco, B. Weng, K. E. Barner, ECG signal denoising and baseline wander correction based on the empirical mode decomposition, Computers in Biology and Medicine 38 (1) (2008) 1–13.

[15] K. Kærgaard, S. H. Jensen, S. Puthusserypady, A comprehensive performance analysis of EEMD-BLMS and DWT-NN hybrid algorithms for ECG denoising, Biomedical Signal Processing and Control 25 (2016) 178–187.

[16] H. Hao, H. Wang, N. Rehman, A joint framework for multivariate signal denoising using multivariate empirical mode decomposition, Signal Processing 135 (2017) 263–273.

[17] P. Van Overschee, B. De Moor, A unifying theorem for three subspace system identification algorithms, Automatica 31 (12) (1995) 1853–1864.

[18] G. Mercère, L. Bako, S. Lecoeuche, Propagator-based methods for recursive subspace model identification, Signal Processing 88 (2008) 468–491.

[19] G.-A. Hossein-Zadeh, B. A. Ardekani, H. Soltanian-Zadeh, A signal subspace approach for modeling the hemodynamic response function in fMRI, Magnetic Resonance Imaging 21 (8) (2003) 835–843.

[20] L. Stanković, D. Mandić, M. Daković, M. Brajović, Time-frequency decomposition of multivariate multicomponent signals, Signal Processing 142 (2018) 468–479.

[21] A. T. Johansson, P. R. White, Instantaneous frequency estimation at low signal-to-noise ratios using time-varying notch filters, Signal Processing 88 (5) (2008) 1271–1288.

[22] J. Liang, B. Ji, J. Zhang, S. Wang, F. Zhao, Recursive least squares-like algorithms for the adaptive second-order lattice notch filter, Digital Signal Processing: A Review Journal 18 (3) (2008) 291–306.

[23] R. Punchalard, A. Lorsawatsiri, W. Loetwassana, J. Koseeyaporn, P. Wardkein, A. Roeksabutr, Direct frequency estimation based adaptive algorithm for a second-order adaptive FIR notch filter, Signal Processing 88 (2) (2008) 315–325.

[24] A. K. Ziarani, A. Konrad, A method of extraction of nonstationary sinusoids, Signal Processing 84 (8) (2004) 1323–1346.

[25] M. Karimi-Ghartemani, A. K. Ziarani, A Nonlinear Time-Frequency Analysis Method, IEEE Transactions on Signal Processing 52 (6) (2004) 1585–1595.

[26] M. Jabloun, N. Martin, F. Leonard, M. Vieira, Estimation of the instantaneous amplitude and frequency of non-stationary short-time signals, Signal Processing 88 (7) (2008) 1636–1655.

[27] B. F. La Scala, R. R. Bitmead, Design of an extended Kalman filter frequency tracker, IEEE Transactions on Signal Processing 44 (3) (1996) 739–742.

[28] S. Bittanti, S. M. Savaresi, On the parametrization and design of an Extended Kalman Filter frequency tracker, IEEE Transactions on Automatic Control 45 (9) (2000) 1718–1724.

[29] R. Kazemi, A. Farsi, M. Ghaed, M. Karimi-Ghartemani, Detection and extraction of periodic noises in audio and biomedical signals using kalman filter, Signal Processing 88 (2008) 2114–2121.

[30] O. Cardona-Morales, L. D. Avendaño-Valencia, G. Castellanos-Domínguez, Nonlinear model for condition monitoring of non-stationary vibration signals in ship driveline application, Mechanical Systems and Signal Processing 44 (1-2) (2014) 134–148.

[31] S. Kim, L. Holmstrom, J. McNames, Multiharmonic tracking using marginalized particle filters., in: Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference, Vol. 2008, 2008, pp. 29–33.

[32] S. Kim, A. S. Paul, E. A. Wan, J. McNames, Multiharmonic frequency tracking method using the sigma-point Kalman smoother, Eurasip Journal on Advances in Signal Processing 2010.

[33] M. West, R. Prado, A. D. Krystal, Evaluation and Comparison of EEG Traces : Latent Structure in Nonstationary Time Series, Journal of the American Statistical Association 94 (448) (1999) 1083–1095.

[34] J. Nakajima, M. West, Dynamic network signal processing using latent threshold models, Digital Signal Processing 47 (2015) 5–16.

[35] H. Sandberg, E. Mollerstedt, B. Bernhardsson, Frequency domain analysis of linear time-periodic systems, IEEE Transactions on Automatic Control 12 (50) (2005) 1971–1983.

[36] E. Louarroudi, J. Lataire, R. Pintelon, P. Janssens, J. Swevers, Frequency domain, parametric estimation of the evolution of the time-varying dynamics of periodically time-varying systems from noisy inputoutput observations, Mechanical Systems and Signal Processing 47 (1-2) (2014) 151–174.

[37] C.-T. Chen, Linear System Theory and Design, 3rd Edition, Oxford University Press, 1999.

[38] M. Verhaegen, X. Yu, A class of subspace model identification algorithms to identify periodically and arbitrarily time-varying systems, Automatica 31 (2) (1995) 201–216.

[39] K. Liu, Identification of Linear Time-Varying Systems, Journal of Sound and Vibration 206 (4) (1997) 487–505.

[40] B. Boashash, Estimating and interpreting the instantaneous frequency of a signal – Part 2: Algorithms and Applications, Proceedings of the IEEE 80 (4) (1992) 540–568.

[41] E. A. Wan, A. T. Nelson, Dual extended kalman filter methods, in: S. Haykin (Ed.), Kalman filtering and neural networks, Wiley, 2001, pp. 123–173.

[42] L. Duque Muñoz, L. D. Avendaño-Valencia, G. Castellanos-Domínguez, Epileptic Seizure Detection in EEG using Time Variant Autoregressive Models, in: A. Jobbagy (Ed.), 5th European IFMBE Conference, Springer, 2011, pp. 1–4.

[43] L. Avendaño-Valencia, S. Fassois, Generalized stochastic constraint time-dependent ARMA modeling of non-stationary signals: Maximum likelihood identification methods, Submitted for Publication on Mechanical Systems and Signal Processing.

[44] G. Kitagawa, Introduction to time series modeling, CRC Press, 2010.

[45] R. H. Shumway, D. S. Stoffer, Time series analysis and its applications, Springer, 2011.

[46] P. Laguna, R. Mark, A. Goldberger, G. Moody, A database for evaluation of algorithms for measurement of QT and other waveform intervals in the ECG, in: Computers in Cardiology, Vol. 24, 1997, pp. 673–676.

[47] J. M. Łęski, N. Henzel, ECG baseline wander and powerline interference reduction using nonlinear filter bank, Signal Processing 85 (2005) 781–793.

[48] S. Agrawal, A. Gupta, Fractal and EMD based removal of baseline wander and powerline interference from ECG signals, Computers in Biology and Medicine 43 (2013) 1889–1899.

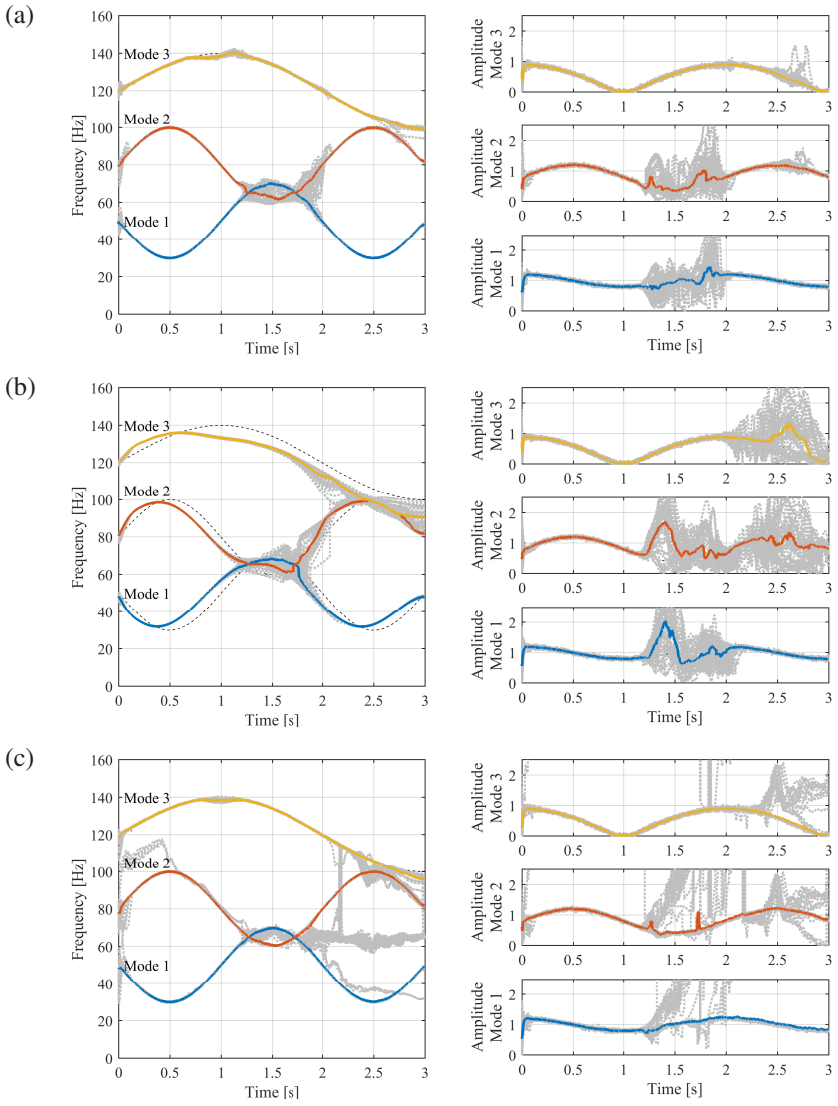


Figure 2: Median and sample-based distribution of the IF and IA estimates obtained with (a) the JKF-NST method, (b) the DKF-NST method, (c) the EKF-FT method. Results obtained after 40 Monte Carlo runs with SNR = 20 dB. Continuous lines represent the estimated mean value, gray dotted lines represent individual sample estimates, black dashed lines represent the actual IF and IA values.

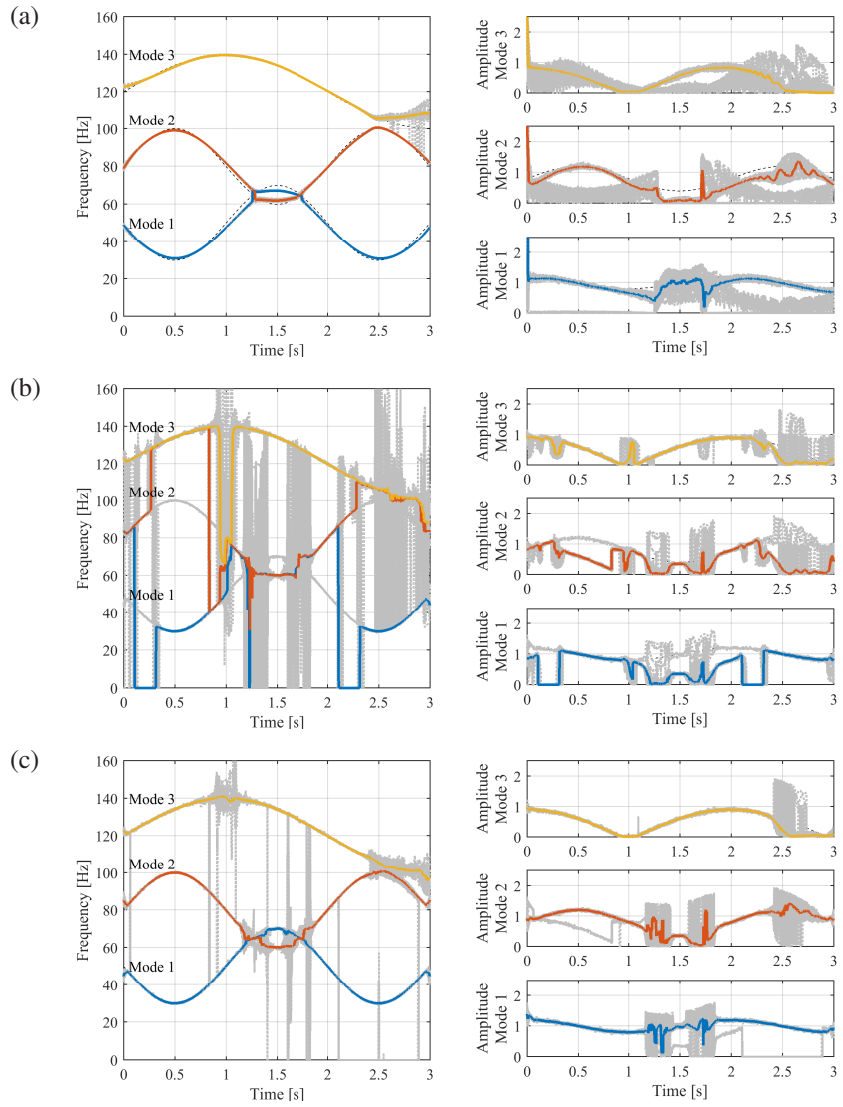


Figure 3: Median and sample-based distribution of the IF and IA estimates obtained with (a) the TAR-MD method based on TAR(18) model, (b) the TFR-maximum method, (c) the TFR-adaptive method. Results obtained after 40 Monte Carlo runs with SNR = 20 dB. Continuous lines represent the estimated mean value, gray dotted lines represent individual sample estimates, black dashed lines represent the actual IF and IA values.

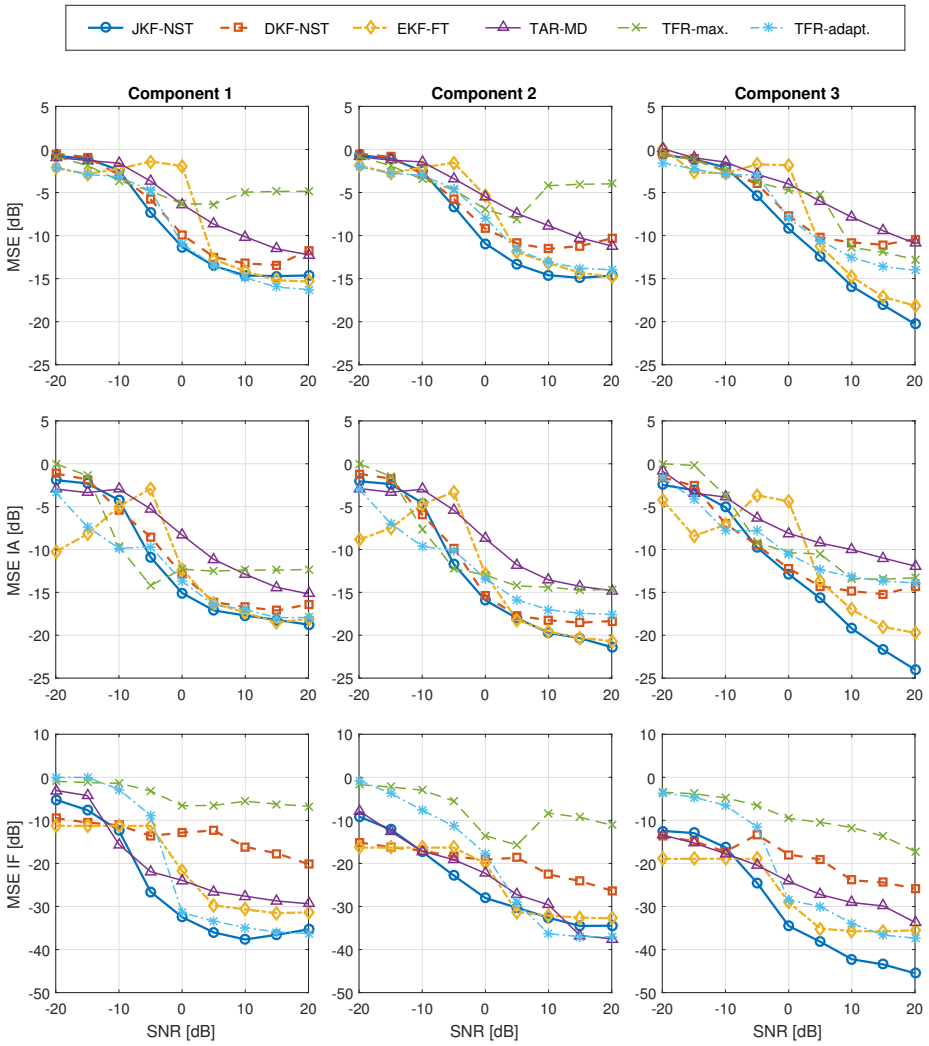


Figure 4: Median normalized MSE of the modal component estimates and their respective IA and IF obtained with the different modal decomposition methods for increasing SNR levels after 40 Monte Carlo runs. Results segregated per component (columns) and according error type (rows). Top row: MSE of the modal component estimates; middle row: MSE of the IA estimates; bottom row: MSE of the IF estimates.

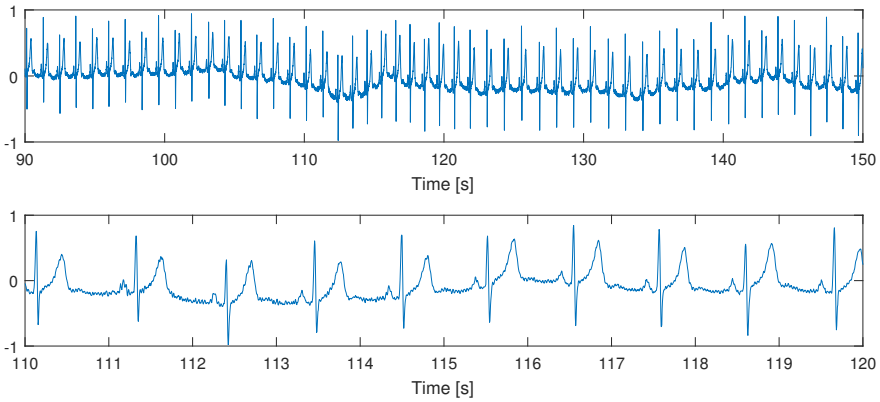
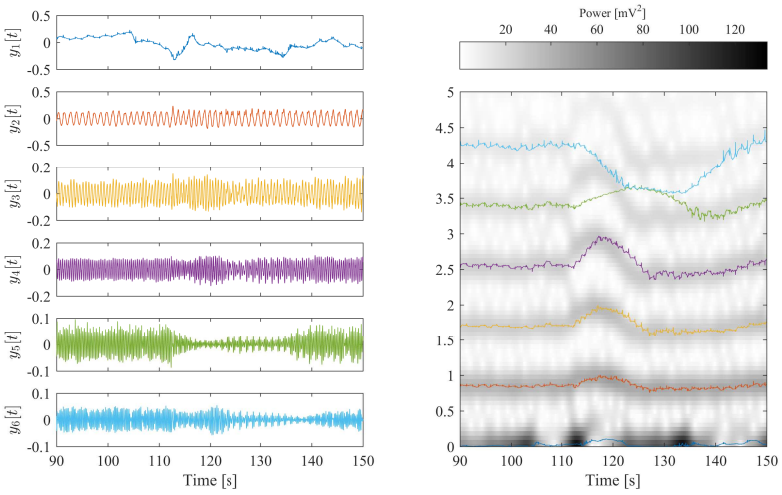
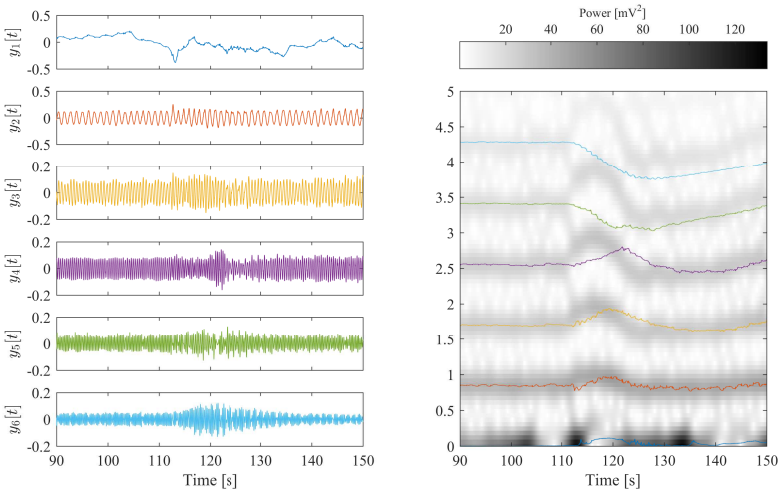


Figure 5: Excerpt from the ECG record "sel117" from the QT database. Top: 60 second portion around $t = 120$ s; Bottom: portion from $t = 110$ s to 120 s.

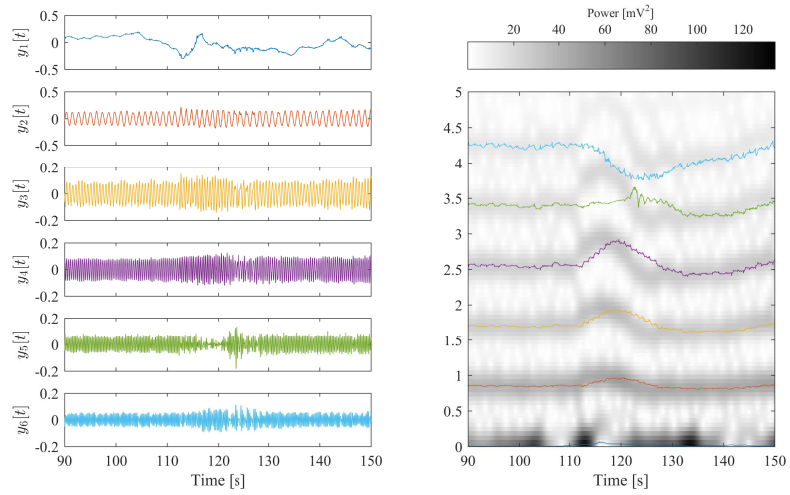


(a)

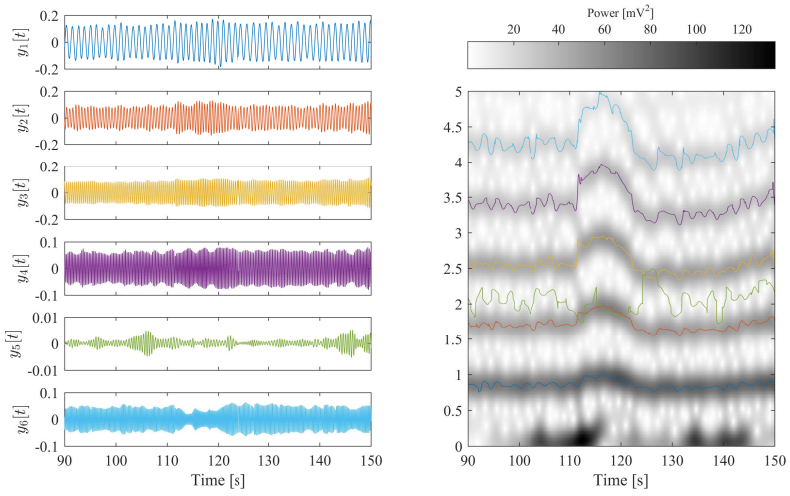


(b)

Figure 6: Modal decompositions of the ECG record "sel117" obtained with (a) the JKF-NST method (b) the DKF-NST method. Left column shows single modal trajectories, while the right column shows the respective modal frequencies laid over the spectrogram of the ECG signal. Spectrogram computed using a Gaussian window of 4096 samples and 4094 sample overlap.



(a)



(b)

Figure 7: Modal decompositions of the ECG record "sel117" obtained with (a) the EKF-FT method (b) the TFR-adaptive method. Left column shows single modal trajectories, while the right column shows the respective modal frequencies laid over the spectrogram of the ECG signal. Spectrogram computed using a Gaussian window of 4096 samples and 4094 sample overlap.

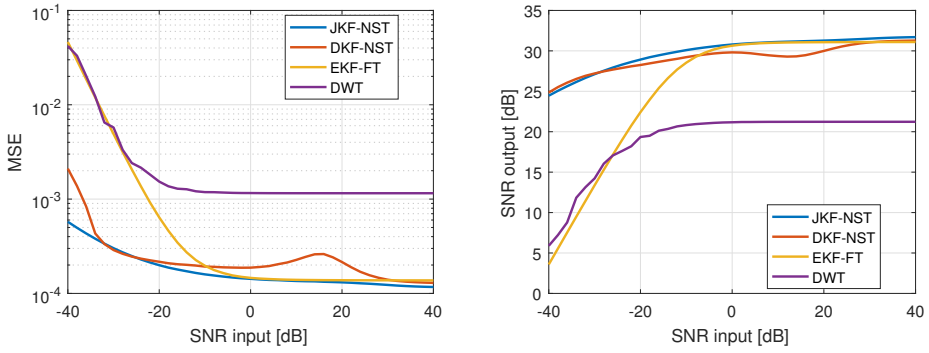


Figure 8: Median SNR and MSE values obtained after removing the powerline interference with the considered denoising methods in the complete set of ECG records of the QT database.