UNIVERSIDAD NACIONAL DE COLOMBIA

Detection of non-stationary dynamics using sub-space based representations, cyclic based and variability constraints

Detección de dinámicas no-estacionarias empleando representaciones basadas en sub-espacios, restricciones de variabilidad y ciclícas

Cristian Castro Hoyos

Universidad Nacional de Colombia
Faculty of Engineering and Architecture
Department of Electrical, Electronic and Computer Engineering, Manizales, colombia
2013

Detection of non-stationary dynamics using sub-space based representations, cyclic based and variability constraints

Detección de dinámicas no-estacionarias empleando representaciones basadas en sub-espacios, restricciones de variabilidad y ciclícas

Cristian Castro Hoyos

Thesis submitted as partial requirement to apply for the degree of:
Master of Engineering - Industrial Automation

Advisor:
Ph.D. Germán Castellanos Domínguez

Research Area:
Digital Signal Processing
Research Group:
Signal Processing and Recognition Research Group

Universidad Nacional de Colombia
Faculty of Engineering and Architecture
Department of Electrical, Electronic and Computer Engineering, Manizales, colombia
2013

Dedicated to...

To my mother and my Grandmother they have been my inspiration and support to follow my dreams.
To my brother, as he has given me the support and pushed me to reach the highest goals.
To my friends, always there, teaching me how to be better.

Some men see things as they are and ask why. Others dream things that never were and ask why not. - *George Bernard Shaw*

Acknowledgements

To anyone and everyone who directly or indirectly was part in the development of this work. First of all, to the professor Germán Castellanos, who lead this research. To all the members of the Signal Processing and Recognition Research group, without their support and collaboration this would not have been possible.

I liked to thank specially to my friend and mentor Diego Hernán Peluffo, he take me upon this path and has always been there for me, for advise, reprimand and joke. To Juan David Martínez, as the last year he was the one who took me under his wing and became my friend.

To David Cárdenas, Andrés Marino, Lina María Sepulveda and Oscar Cardona, who have been great and dear colleagues always helping me, it is from seen people like you that one can see research and science its not a work, it is just a hobby that you get paid for.

Additionally, i want to thank to the institutions that made this happen, the program *Jovenes Investigadores e Innovadores 2011*, "Virginia Gutierrez Pineda", financed by COLCIENCIAS, the Researh Direction of Universidad Nacional de Colombia, Manizales, DIMA and finally the project *Grupo de Control y Procesamiento Digital de Señales, Código: 20501007502*.

Finally, to my family, my mother, my grandmother and my brother, they continuously supported me, always letting me pursuit my dream putting on hold their own. Hopefully some day, i will help them to reach them.

To all of you and those who i might forgotten, thanks. Let this to be just a stair in a path full of achievement to enjoy together.

Cristian Castro Hoyos

Resumen

La siguiente Tesis de Maestría propone una metodología para el análisis de series de tiempo no-estacionarias con el propósito de filtrado y detección de ruido en reconocimiento de patrones. La metodología se encuentra dividida en dos etapas: el análisis de comportamientos no-estacionarios que recaen en procesos cíclicos y como diferentes componentes no-periódicos afectan el análisis de la señal. El segundo enfoque, está centrado en el problema de extracción de series de tiempo no-estacionarias que afectan procesos estacionarios. Ambos esquemas están basados en restricciones de (ciclo-)estacionariead y representaciones basadas en sub-espacios de manera que mediante la evaluación de las dinámicas de la señal sea posible identificar las componentes no-estacionarias indeseadas. Los resultados se muestran para cada enfoque de manera independiente por medio de datos sintéticos y reales, el desempeño obtenido muestra una gran capacidad de detección, rechazo y/o extracción de ruido y artefactos en series de tiempo (ciclo-)estacionarias usando restricciones de estacionariedad así como condiciones cíclicas basadas en la naturaleza de la señal.

Palabras claves: Ciclo-estacionariedad, Non–estacionario, Estacionario, Representación Basada en Sub-espacios, Filtrado, Representación Tiempo-Frecuencia.

Abstract

The present Master's Thesis proposes a methodology for the non–stationary time-series analysis for filtering and noise rejection purposes in pattern recognition. The methodology is divided into two different approaches: the analysis of periodic non–stationary behavior that relies into a cyclic process and how additional non–cyclic non–stationarities disrupt and affect the signal processing. Second approach deals with the problem of non–stationary signal extraction that affects inherent weak stationary processes. Both frameworks of analysis take base on (cyclo-)stationary constraints and subspace based representations in order to assess and characterize the signals dynamics to facilitate the identification of the undesired non–stationary components. Results are shown over each approach with different real and synthetic data, the obtained performances show high rejection, detection and extraction capabilities for noise and artifacts in (cyclo)–stationary signals using external and internal based constraints of analysis and high separation capability for stationary signals.

Keywords: Cyclo-Stationary, Non–Stationary, Stationary, Subpspace Based representation, filtering, time-frequency representation.

Contents

List of Figures

List of Tables

List of Algorithms

Notation

Variables and Functions

i, j, k	Subindex
$\boldsymbol{X}$	Data Matrix
$\boldsymbol{x}$	Data Vector
$y\left(t\right)$	Time Series
Θ	Statistic Estimator
$\epsilon, \eta.$	Thresholds
μ	Mean
σ	Standard Deviation
τ, ζ	Time Window
α	Cycle Frequency
$R_y\left(t, \tau\right)$	Autocorrelation
R_y^α	Cyclic Autocorrelation
$S_y\left(t, f\right)$	Cycle Frequency Spectrum
t^-	Previous Time instant
$\delta\left(t\right)$	Instantaneous Cycle frequency
$n.$	Signature Number of frames
e_y^i	Energy
$e_y\left(t\right)$	Energy Envelope
$\tilde{e}$	Average Energy
$\rho_y, \theta_y, \psi_p, \chi^j$	Cyclic Signatures
λ	Eigenvalues
$\kappa., \beta., c.$	Function Parameters
p_j^i	Peak
s_e	Sensitivity
$\boldsymbol{\Phi}$	Subspace Matrix
$y_s\left(t\right)$	Stationary Signal
$y_n\left(t\right)$	Non–Stationary Signal
$\boldsymbol{\nu}$	Set of eigenvalues
$\boldsymbol{\Sigma}$	Covariance Matrix

Mathematical Operators

$\boldsymbol{E}\{\cdot\}$	Expected Value
$\mathcal{N}(\mu, \boldsymbol{\Sigma})$	Gaussian Process
$\lfloor \cdot \rfloor$	Nearest Integer Approximation
$\langle \cdot \rangle$	Inner Product
$\|\cdot\|_L$	L-norm
$\{\cdot\}$	Set of values
$\mathcal{M}\{\boldsymbol{x}, \epsilon\}$	Separation Operator
$var\{\cdot\}$	Variance
$\mathrm{vec}(\cdot)$	Vectorization Operator
$\mathrm{Diag}\cdot$	Diagonal matrix of its argument
$cov\{\cdot,\}$	Covariance
$\mathrm{d}(\cdot)$	Distance or Similarity measure

Abbreviations

EEG	Electroencephalogram
PCG	Phonocardiogram
BSS	Blind Source Separation
BCE	Blind Component Extraction
WSS	Wide Sense Stationary
STCFS	Short Time Cycle Frequency Spectrum
TFR	Time Frequency Representation
ICF	Instantaneous Cycle Frequency
HRV	Heart Rate Variability
CWT	Continuous Wavelet Transform
SNR	Signal-to-Noise Ratio
ASSA	Analytic Stationary Subspace Analysis
SONS	Second Order Non–Stationary Sources
SSA	Singular Spectrum Analysis
PCA	Principal Component Analysis
SS-SC	Subspace-based Separation with Stationarity Constraints
PSWVD	Pseudo-Wigner Ville Distribution
R-SSA	Recursive Singular Spectrum Analysis
RSS-SC	Recursive Subspace-based Separation with Stationarity Constraints
SVD	Singular Value Decomposition
DS	Degree of Stationarity
const	Constant

Part I

Preliminaries

1. Introduction

1.1 Subspace Stationary Analysis for Signal Separation

Separation methods from mixed signals with a priori knowledge are usually required in applications, such as in speech signal processing, biomedical signal analysis, communication, econometrics, and neuroscience. Examples of concrete tasks include: to recover radar signals contaminated with highly moving targets [33] or strong sea-clutter [34], to improve interpretation of geomagnetic measurements aiming at factorizing the observed time series into meaningful components [38], to filter high frequency activity not related to motor tasks in EEG data [18], to develop non uniform embedding reconstruction of pathological voices [17], among others. In those specific contexts, the cornerstone in the filtering task is the separation model holding certain stochastic constraints imposed to make composite signals statistically disjoint. This work concerns separation between non-stationary and stationary signals that inherently assume the stationarity of one of the composite signals. To this end, time variability of probability function densities may be established as a separability criterion, yet, the most common approach is based on more relaxed condition, by instance, second-order statistics invariability along the time, i.e., either the correlation function/power spectrum or the mean and variance values [5]. However, one of the major concerns in using many of separation filtering methods, including those devoted to extract non-stationary signals, is the poor parameter estimation due to influence of non-stationary structure. To avoid the situation, methods tend to include time-variability of stochastic structure within the estimation model. By instance in [7], separating algorithms are developed by jointly tackling the non-stationary and temporally correlated mixing coefficients and source signals through the online Gaussian process. A similar approach is developed in [19], but using continuous density hidden Markov models. Likewise in [30], evolutionary spectral density functions of uniformly modulated and strongly narrow-band stochastic processes are modeled. In [25], assumption is made that the time-frequency distributions of considered input signals do not overlap. Nonetheless, those methods, mostly, do allow separating signals under very specific assumed non-stationary properties. A less restrictive separating approach can be developed grounded in conventional nonparametric subspace analysis that consists of searching such a

projection maximally bearing input information when retaining only those data that contribute most to data stationarity representation. As a result, projection provides signals that are as stationary as possible, while introducing new bias variability measures capable of quantize the signal stochastic stationarity might posed solutions under less restrictive assumptions, such at non-specific number of stationary sources, stochastic local stationarity and convolutive mixture model identification from spectral analysis.

On the other hand, cyclo–stationary signals are a special case of non–stationary processes that repeat them selves with certain periodicity [12], holding similar definitions to those of stationary processes but within an specific time lag or period. In the recent years, an increasing interest of cyclo-stationary signal analysis as been reported [10], from telecommunications to vibration signals [40] and biological signals regarding cardiac function [2]. Cyclo–stationary stochastic properties are commonly analyzed by means of the periodicity in the autocorrelation functions, thus, taking the fourier expansion of such function, referred as cyclic autocorrelation functions, the periodic nature of the process becomes more visible. Purely time domain analysis of cyclo–stationary signals is quite complex, since conventional spectral analysis has been design for stationary signals some problems arises when directly applied over cyclo–stationary process, such as the concept of two dimensional spectrum. However, as cyclo–stationary signals are usually local non–stationary, a direct extraction or separation approach can not be implemented, instead, there is a primordial need for identification of the cyclic properties of the phenomena, such as cycle frequency estimation, spectral coherence, temporal cycle similarity, etc. prior to any separation task[3].

2. Problem Statement

Classic digital signal processing takes upon a common assumption the general stationarity of the studied signals, restriction under which the signal stochastic properties are time-invariant. However, biological signals, mechanic signals and some other processes, require the direct study of the non–stationary components as main source of information as the time-variant stochastic features are directly related to the underlying studied phenomena[41]. Specifically, non–stationary processes can affect not only assumed narrow sense stationary signals, but also non–stationary signals that present with a periodic behavior, as some other non–stationary components not. This cyclo–stationary processes are usually analyzed via second order statistics and the component extraction of underlying stationarity, cyclo–stationarity and non–stationarity is done by blind component extraction methodologies that work under specific constraints and makes use of a priori given information. However, when the spectrum of the non–stationary cyclic process overlaps with the non–periodic non–stationary components spectrum such separation its not an easy task. Moreover, the presence of such non–stationary corrupting signals might result in inaccurate extracted components and by consequence any other posterior processing would be compromised. Given this facts, the unsupervised identification of non–stationary components over cyclo–stationary signals arises as a first stage of preprocessing task, where based only in external based factors, the cyclo–stationary features of the signals, a quality criteria over the signal cyclo–stationarity information is given, aiming to provide information to the consequent processing analysis in order to either extract the non–stationarities from the cyclic process or reject the signal for analysis given the low quality and lack of information.

In addition to the cyclo–stationarity analysis, second order statistic are also useful to provide a signal analysis framework where non–stationary processes are to be separated from underlying stationary signal analysis. However, common blind source separation methodologies require of specific a priori restrictions as multi-dimensional input data, number of stationary sources, locally stationarity and others [22],[20]. Upon this facts, a one dimensional time–series stationarity signal analysis becomes necessary, as in a Blind Component extraction process, where given the studied dynamics of the input data, it becomes possible to identify underlying stationary and non–stationary process from enhanced representation spaces.

Generally speaking, used in separation between non-stationary and stationary signals, subspace-based methods should have as desired properties the following: *i*) to highlight stochastic behavior of an underlying random process during projection [26], *ii*) to ensure established stochastic constraints dealing with non-stationary signal structure within the projection framework. As regards the former ability, in [42], several projection methods are considered for extracting information from non-stationary time series, concluding that a bundle of methods may lead to much extracted information about the process. But, to match the stochastic constraints after such a complex transformation may pose a challenge. Rather, to enhance the time-series representation within the subspace framework, any of the standard methods for analysis of non-stationary time series can be suitable. In case of the latter ability, a separation method should be carried out imposing the stochastic constraints upon the linearly transformed space. However, projection of non–stationary signals is not always identifiable. Particularly, the most widely used constraint requires that the second-order statistics time invariability must coincide with the ensemble averages, which is far from being solved even for linear subspace models. Therefore, to discriminate between stationarity and non-stationarity, we must still evaluate how stochastic restriction holds over the optimization solution within a given projection framework.

3. Objectives

3.1 General Objective

Develop a feature representation framework based on subspace projections and (cyclo-) stationarity constraints in order to enhance the detection and/or extraction capabilities of non-stationary dynamics in time-series.

3.2 Specific Objectives

1. Develop a time-frequency based representation framework that assess the presence of non-stationary components in biological time-series, using adaptive temporal and spectral signatures that highlight signal cyclo-stationary features.

2. Develop a subspace based filtering methodology to extract the non–stationary components in a given time–series using weakly stationary constraints and stochastic feature relevance analysis. The proposed scheme would be tested over synthetic and real time–series in order to enhance results interpretability, this is assessed by means of non–stationarity measures and spectral analysis of filtered components.

3. Design a time–variant feature representation framework based on subspace signal projections to assess the signal stationarity behavior under a dynamic recursive environment. To this, new collected data would update the signal subspace structure taking into consideration the new information and its relation with previous collected data.

Part II

Detection of Non-stationary Components using adaptive cyclo-stationary signal analysis

In this chapter, common stationary characteristics of signal processing are taken into consideration for analysis using enhanced time–frequency representations. The formulation of stationary and cyclic stationary signals statistics constraints and how such parameters are to be used in posterior processing tasks related to the signal pattern recognition processes is first given. As the cyclo–stationary signal behavior is directly related to the generating process, several additional constraints are imposed to generate a more adaptive and interpretable analysis. Finally, results are presented using common cyclo–stationary signals, as those related to cardiological functions. It is highlighted how primary detection of interference components can affect posterior signal processes as segmentation, feature extraction and/or classification.

Mathematically, the proposed framework is presented in order to provide a methodology for cyclo–stationary signal analysis aimed to detect components that are not related within the signal process information. Generally speaking, cyclo–stationary signals are regarded into the wide–band spectrum signals, thereby signal filtering is not always a possibility and only signal rejection or unreliable result alerts are given as options. Blind component extraction of cyclo-stationary signals, classification, segmentation and any other further signal processing are out of the scope of the present master thesis and thus only basic procedures as used for performance analysis.

4. Signal Cyclo–Stationary Analysis from Second Order Statistics

4.1 Signal cyclo-stationarity definition

Stationarity in the narrow sense, requires the constant behavior of a given set of statistical moments of a time-series process $y(t) \in \mathbb{R}$ over a fixed period of time,

$$\lim_{\tau \to T} \boldsymbol{E}\{|\Theta\{y(t)\} - \Theta\{y(t+\tau)\} : \forall \tau \in T|\} - \epsilon \leq 0 \, , \forall t \in T \tag{4-1}$$

where $\Theta\{\cdot\}$ stands for an statistic value estimator and $\epsilon \subset \mathbb{R}^+$ is a certain small enough value that express the degree of accuracy of the given stationarity.

Any given time-series, that with an appropriate values of time shifting window τ and degree of accuracy ϵ fails to convey this constraint is assumed to be non-stationary or more accurately, does not fulfill weak-sense stationarity. Generally speaking it can be assumed that weak stationarity is measurable with the first two moments, mean and variance, noted respectively as $\Theta \in \left\{\mu_{\boldsymbol{y}}, \sigma^2_{\boldsymbol{y}}\right\}$.

Additionally to strictly stationary and non-stationary processes, a given stochastic process whose statistical moment values vary periodically with time [14], generally speaking, a non–stationary process that repeats itself over fixed windows of time, are classified as cyclo–stationary processes which arises as the following definition:

Definition 4.1.1. (*Cyclo–Stationary Process*): Let $y(t)$ be a random signal trajectory, it is said to be a cyclo–stationary process in the wide sense, with cycle period τ_c, if the following cyclo-stationary constraints holds,

$$\boldsymbol{E}\{y(t)\} = \boldsymbol{E}\{y(t+\tau_c)\} \tag{4-2a}$$

$$R_{\boldsymbol{y}}(t, \tau_c) = R_{\boldsymbol{y}}\left(t^{'}, \tau_c\right) \tag{4-2b}$$

with $R_{\boldsymbol{y}}(t,\tau_c)$ is the correlation function of the time series $\boldsymbol{y}$ defined as,

$$\boldsymbol{R}_{\boldsymbol{y}}(t,\tau_c) \triangleq \frac{\boldsymbol{E}\{(y(t)+\mu_{\boldsymbol{y}})(y(t+\tau_c)-\mu_{\boldsymbol{y}})\}}{\sigma_{\boldsymbol{y}}^2} \tag{4-3}$$

Consequently, the definition 4.1.1 can be posed into a more relaxed form, as noted by [12], where a process is said to be *almost cyclo-stationary in the wide sense* if there exists a periodicity frequency $\alpha = {}^{1}\!/_{\tau_c}$ for which the Fourier series expansion coefficients of the cyclic autocorrelation function are not identically zero [12].

4.2 Cyclic Autocorrelation and Cyclic Frequency Spectrum

Given the Fourier series expansion of the autocorrelation function of a non–stationary process,

$$\boldsymbol{R}_{\boldsymbol{y}}(t,\tau_c) = \sum_{\alpha} R_{\boldsymbol{y}}^{\alpha}(t,\tau_c)\, e^{-j2\pi\alpha t} \tag{4-4a}$$

$$\boldsymbol{R}_{\boldsymbol{y}}^{\alpha}(t,\tau_c) \triangleq \lim_{T\to\infty}\frac{1}{T}\int_{-\frac{T}{2}}^{\frac{T}{2}} R_{\boldsymbol{y}}(t,\tau_c)\, e^{-j2\pi\alpha t}dt \tag{4-4b}$$

for a value of cycle frequency $\alpha = 0$, results in the time-averaged probabilistic autocorrelation function for asymptotically mean stationary processes, which indicates that the cyclic autocorrelation function is time-invariant for $\alpha = 0$, however this same condition does not hold for $\alpha \neq 0$, rather they are cyclicly repeated as in a sinusoidal modulated function, that is,

$$R_{\boldsymbol{y}}^{\alpha}(t+t_0,\tau_c) = R_{\boldsymbol{y}}^{\alpha}(t,\tau_c)\, e^{j2\pi\alpha t_0} \tag{4-5}$$

And the set of frequencies α for which $R_{\boldsymbol{y}}^{\alpha}(t,\tau_c) \not\equiv 0$ is called the cycle spectrum of $y(t)$.
As in common stationary and non–stationary signal analysis, the power spectral density of a given time series $y(t)$ can be defined from the autocorrelation function $R_{\boldsymbol{y}}$ as,

$$S_{\boldsymbol{y}}(f) = \int_{-\infty}^{\infty} R_{\boldsymbol{y}}(t,\tau_c)\, e^{-j2\pi f\tau_c}d\tau_c \tag{4-6}$$

whereas it can be proven that for cyclo–stationary signals, the cyclic autocorrelation of process $y(t)$ is defined as the time cross-correlation of frequency shifted versions of $y(t)$,

which leads to the cyclic spectral density and cyclic spectrum as follows:

$$S_{\boldsymbol{y}}(f) = \int_{-\infty}^{\infty} R_{\boldsymbol{y}}^{\alpha}(t,\tau_c)\, e^{-j2\pi\tau_c f} d\tau_c \triangleq S_{\boldsymbol{y}}^{\alpha} \tag{4-7a}$$

$$S_{\boldsymbol{y}}(t,f) = \sum_{\alpha} S_{\boldsymbol{y}}^{\alpha}(f)\, e^{j2\pi\alpha t} \tag{4-7b}$$

Broadly, cyclo–stationary processes are characterized by the existence of spectral correlation within the process and this is parameterized by the cyclic spectral density $\left\{S_{\boldsymbol{y}}^{\alpha}\right\}$ or by the cyclic autocorrelations set $\left\{R_{\boldsymbol{y}}^{\alpha}\right\}$.

4.3 Time-Frequency Analysis of Cyclo–stationary Signals

Stationary signal analysis provides with a tool for examining the degree to which two wide sense stationary (WSS) processes are related by time-invariant linear transformations, this measured, called cross–coherence function $\rho(f)$ is defined by:

$$\rho(f) = \frac{S_{\boldsymbol{xy}}(f)}{\left[S_{\boldsymbol{x}}(f)\, S_{\boldsymbol{y}}(f)\right]^{1/2}} \tag{4-8}$$

The same concept might be applied to *almost cyclo–stationary signals* where the coherence function of frequency shifted versions of the process $y(t)$ exhibits cyclo–stationarity if and only if $\rho_{\boldsymbol{y}}^{\alpha}(f)$ exists, that is,

$$\rho_{\boldsymbol{y}}^{\alpha}(f) = \frac{S_{\boldsymbol{y}}^{\alpha}(f)}{\left[S_{\boldsymbol{y}}^{\alpha}(f+\alpha/2)\, S_{\boldsymbol{y}}^{\alpha}(f-\alpha/2)\right]^{1/2}} \tag{4-9}$$

Equation 4-9 is called the auto-coherence spectral function and it indicates the degree of spectral coherence of the process $y(t)$ at spectral frequency f and cycle frequency α, with $|\rho_{\boldsymbol{y}}^{\alpha}(f)| = 1$ for completely coherent processes and $|\rho_{\boldsymbol{y}}^{\alpha}(f)| = 0$ for completely incoherent processes at spectral frequency f and cycle frequency α.

However, as the stationary process, any cyclo–stationary process might be affected by any other non–stationary components that do not have spectral correlation at given cycle frequencies, this means that cycle frequency can vary with time within a limited range of the fundamental cycle frequency α, this introduces then, the necessity of a time-varying estimation of the cyclic autocorrelation function and by extension the cycle frequency spectrum and cyclic auto–coherence function. [35] introduced the short-time frequency spectrum by com-

puting the cyclic autocorrelation function in predefined sliding windows in the time domain,

$$R_{\boldsymbol{y}}^{\alpha}(t,\tau) = \int_{t-\zeta}^{t+\zeta} R_{\boldsymbol{y}}(t,\tau)\, e^{-j2\pi\alpha t} dt \tag{4-10a}$$

$$S_{\boldsymbol{y}}^{\alpha}(t,f) = \int_{-\infty}^{\infty} R_{\boldsymbol{y}}^{\alpha}(t,\tau)\, e^{-j2\pi f\tau} d\tau \tag{4-10b}$$

$$\rho_{\boldsymbol{y}}^{\alpha}(t,f) = \frac{S_{\boldsymbol{y}}^{\alpha}(t,f)}{\left[S_{\boldsymbol{y}}^{\alpha}(t,f+{}^{\alpha}\!/_{2})\, S_{\boldsymbol{y}}^{\alpha}(t,f-{}^{\alpha}\!/_{2})\right]^{1/2}} \tag{4-10c}$$

It can be seen in equation (4-10c) that the set of coefficients of the cyclic auto–coherence function $\{\rho_{\boldsymbol{y}^{\alpha}}(t,f)\}$ covers three different domains, spectral frequency, cyclic frequency and time, yet we can make this a function of time and cycle frequency by integrating the function over the frequency spectrum to obtain the short–time cycle frequency spectrum (STCFS) as

$$\gamma_{\boldsymbol{y}}(t,\alpha) = \int_{-\infty}^{+\infty} \rho_{\boldsymbol{y}}^{\alpha}(t,f)\, df \tag{4-11}$$

It should be remarked that the time window size for computing the STCFS covers $[t-\zeta, t+\zeta]$ and a necessary condition for the cycle frequency spectrum to exist over this interval is imposed as at least two cycles of the periodic non–stationary process must occur over this interval [12].

In [36] it is pointed out that the time lag parameter of equation (4-10a) can also be dropped or more accurately set to $\tau = 0$, since the autocorrelation function at cycle frequency of interest must be approximately equal to those of lagged versions of the signal. Thereby equations (4-10a) to (4-11) can be written as,

$$R_{\boldsymbol{y}}^{\alpha}(t) = \int_{t-\zeta}^{t+\zeta} y(t)^2\, e^{-j2\pi\alpha t} dt \tag{4-12a}$$

$$\gamma_{\boldsymbol{y}}(t,\alpha) = \frac{R_{\boldsymbol{y}}^{\alpha}(t)}{R_{\boldsymbol{y}}^{0}(t)} \tag{4-12b}$$

The short time frequency spectrum of equation (4-11) holds then the cycle frequency information of the process $y(t)$ in a cycle-frequency time map representation. Thereby, in [35] it is propose the extraction of the Instantaneous Cycle Frequency (ICF) $\delta_{\boldsymbol{y}}(t)$ as the cycle

frequency band greater that zero to which $\gamma_{\boldsymbol{y}}(t,\alpha)$ presents the higher energy concentration,

$$\delta_{\boldsymbol{y}}(t) = \arg\max_{\alpha} \left[\gamma_{\boldsymbol{y}}(t,\alpha)\right]_{\delta_{\boldsymbol{y}}(t^-)-\lambda_\delta}^{\delta_{\boldsymbol{y}}(t^-)+\lambda_\delta} \forall t \in \mathbb{R}^T \tag{4-13}$$

where $\delta_{\boldsymbol{y}}(t^-)$ indicates the previous value of the estimated ICF and λ_δ is a cycle frequency boundary that controls the variational limits of the newer $\delta_{\boldsymbol{y}}(t)$ to a maximum deviation of λ_δ.

4.4 Fast STCFS Estimation

General computation of the short cycle-frequency spectrum with sliding window as in a short time fourier transform directly trough (4-12a) to (4-12b) is computationally too expensive [13], however when a defined cycle frequency spectrum is stated as $\boldsymbol{\alpha} \in [\alpha_l, \alpha_h]$ the direct computation of the STCFS can be achieve by means of recursive operations and inner products among lagged signal samples and fourier basis of the cycle frequency.

Let $y(t) \in \mathbb{R}^T$ be a non–stationary process with periodicity over the cycle frequency $\boldsymbol{\alpha} \in [\alpha_l, \alpha_h]$ and let $\zeta_{\boldsymbol{y}}$ be a time window such that 3 complete periods of the non–stationary behavior are within such window. The unnormalized STCFS of equation (4-12a) for a time sample point t is given by,

$$\boldsymbol{R}_{\boldsymbol{y}} = \begin{bmatrix} \sum_{k=1}^{1+2\zeta_{\boldsymbol{y}}} y[k]\, e^{-j2\pi\alpha_l k} & \sum_{k=2}^{2+2\zeta_{\boldsymbol{y}}} y[k]\, e^{-j2\pi\alpha_l k} & \cdots & \sum_{k=T-2\zeta}^{T} y[k]\, e^{-j2\pi\alpha_l k} \\ \sum_{k=1}^{1+2\zeta_{\boldsymbol{y}}} y[k]\, e^{-j2\pi\alpha_{l+1} k} & \sum_{k=2}^{2+2\zeta_{\boldsymbol{y}}} y[k]\, e^{-j2\pi\alpha_{l+1} k} & \ddots & \vdots \\ \sum_{k=1}^{1+2\zeta_{\boldsymbol{y}}} y[k]\, e^{-j2\pi\alpha_h k} & \sum_{k=2}^{2+2\zeta_{\boldsymbol{y}}} y[k]\, e^{-j2\pi\alpha_h k} & \cdots & \sum_{k=T-2\zeta}^{T} y[k]\, e^{-j2\pi\alpha_h k} \end{bmatrix}_{f\times T} \tag{4-14}$$

Equation (4-14) increase its computational burden complexity as the length of the signal of the cycle frequency range increases, however it can be seen that the summations required over each column of $\boldsymbol{R}_{\boldsymbol{y}}$, can be directly computed as the inner product of the window vector

of the signal $\boldsymbol{y}$ and a matrix of Fourier basis functions of $\boldsymbol{\alpha} \in \mathbb{R}^f$ as follows

$$\boldsymbol{B}_a^b = \left[\begin{array}{cccc} \left(e^{-j2\pi a\boldsymbol{\alpha}}\right)^{\mathsf{T}} & \left(e^{-j2\pi a+1\boldsymbol{\alpha}}\right)^{\mathsf{T}} & \cdots & \left(e^{-j2\pi b\boldsymbol{\alpha}}\right)^{\mathsf{T}} \end{array} \right]_{f\times 2\zeta} \tag{4-15a}$$

$$\boldsymbol{R}_{\boldsymbol{y}} = \left[\begin{array}{cccc} \left(\boldsymbol{y}_1^{1+2\zeta} \cdot \boldsymbol{B}_1^{1+2\zeta}\right)^{\mathsf{T}} & \left(\boldsymbol{y}_2^{2+2\zeta} \cdot \boldsymbol{B}_2^{2+2\zeta}\right)^{\mathsf{T}} & \cdots & \left(\boldsymbol{y}_{T-2\zeta}^{T} \cdot \boldsymbol{B}_{T-2\zeta}^{T}\right)^{\mathsf{T}} \end{array} \right]_{f\times T} \tag{4-15b}$$

At this point, the complexity of multiple iterative computations over the different values of $\boldsymbol{\alpha}$ has been reduced to a matrix based product, yet this procedure can still be reduced such that for the STCFS at sample point t is obtained by a rank-one update of the previously sample point $t-1$, to demonstrate this fact, lets recall the summation based form of equation (4-14) for a couple of given sample points $t, t+1$

$$\boldsymbol{R}_{\boldsymbol{y}}(t) = \left[\begin{array}{c} \sum_{k=t}^{t+2\zeta} y[k]\, e^{-j2\pi\alpha_l k} \\ \sum_{k=t}^{t+2\zeta} y[k]\, e^{-j2\pi\alpha_{l+1} k} \\ \cdots \\ \sum_{k=t}^{t+2\zeta} y[k]\, e^{-j2\pi\alpha_h k} \end{array} \right] \quad \boldsymbol{R}_{\boldsymbol{y}}(t+1) = \left[\begin{array}{c} \sum_{k=t+1}^{t+1+2\zeta} y[k]\, e^{-j2\pi\alpha_l k} \\ \sum_{k=t+1}^{t+1+2\zeta} y[k]\, e^{-j2\pi\alpha_{l+1} k} \\ \cdots \\ \sum_{k=t+1}^{t+1+2\zeta} y[k]\, e^{-j2\pi\alpha_h k} \end{array} \right] \tag{4-16a}$$

$$\boldsymbol{R}_{\boldsymbol{y}}(t+1) = \left[\begin{array}{c} \left(\sum_{k=t}^{t+2\zeta} y[k]\, e^{-j2\pi\alpha_l k}\right) - y[t]\, e^{-j2\pi\alpha_l t} + y[t+2\zeta+1]\, e^{-j2\pi\alpha_l (t+2\zeta+1)} \\ \left(\sum_{k=t}^{t+2\zeta} y[k]\, e^{-j2\pi\alpha_{l+1} k}\right) - y[t]\, e^{-j2\pi\alpha_{l+1} t} + y[t+2\zeta+1]\, e^{-j2\pi\alpha_{l+1} (t+2\zeta+1)} \\ \cdots \\ \left(\sum_{k=t}^{t+2\zeta} y[k]\, e^{-j2\pi\alpha_h k}\right) - y[t]\, e^{-j2\pi\alpha_l t} + y[t+2\zeta+1]\, e^{-j2\pi\alpha_h (t+2\zeta+1)} \end{array} \right] \tag{4-16b}$$

This vector-matrix based form reduces the computational complexity of the STCFS using rank-one updates of the cycle spectrum, such that for a new given sample point there is no need for a matrix inner product but a rank-one update that uses simpler vector sum/rest operations.

4.5 Enhanced Estimation of Instantaneous Cycle Frequency

It should be remarked however, that almost cyclo–stationary processes might hold not only one fundamental cycle frequency, but also some periodic integers harmonics of its periodic nature as well as some other cyclic processes that can corrupt the STCFS representation [36]. Additional the variational drift of the cycle frequency can be greater than the predefined λ_y criteria. Thus, maximum argument of equation (4-13) can produce wrong estimations of ICF locating the cycle frequency over a harmonic value of the fundamental cycle frequency and/or produce an abrupt change of the cycle frequency value if the deviation control parameter is too large and the signal is contaminated with some artifacts. Given this facts, a search algorithm that identify the highest energy band in the STCFS and detects the integer harmonic to which it relates according to the expected cycle frequencies that can be achieved by the cyclo–stationary process is posed, under the only necessary condition that a priori knowledge of the highest harmonic value of the expected cycle frequency α_f is given.
However there are several cases when the true cycle frequency harmonic can not be found and is set to zero:

- One single peak in the autocorrelation function of the α marginal means that the STCFS does not hold any cycle frequency information which can be caused by corrupting noise or badly set boundaries.

- The found prominent peaks are below the expected cycle frequency and cannot be related to cyclo–stationary process.

- The identified harmonic does not achieve minimum fundamental cycle frequency under maximum allowed variation

4.6 Interference Detection in Cyclo-stationary Signals

Frequency spectrum of cyclo–stationary process is not only bounded by the cycle frequency but also by the natures process, generally speaking, filtering or extraction of unwanted components that are affecting the process is required as preprocessing task to any posterior signal analysis, however when signal spectrum is in the same band as the interference spectrum such task is not easy achievable or sometimes not possible [14]. To this, interference and

Algorithm 1 Enhanced ICF Search Algorithm

Input: signal $y(t) \in \mathbb{R}^T$, cycle frequency range of operation $\boldsymbol{\alpha} \in [\alpha_l, \alpha_h]$, initial search deviation λ_δ, ζ window size and maximum harmonic value h_h. **Output:** Corrected ICF $\delta(t)$ and main harmonic detected h.
Compute $\boldsymbol{R_y}(t, \alpha)$ via the Fast STCFS estimation and obtain $\boldsymbol{\gamma_y}(t, \alpha)$. Obtain ICF initial estimate $\delta \boldsymbol{y}^0(t)$. Compute the autocorrelation function of STCFS energy envelope as

$$e_{\boldsymbol{R_y}}(\alpha) = \sum_{\forall t} R_{\boldsymbol{y}}^{\alpha}(t)$$

$$\tilde{R}_{\boldsymbol{e}_R}(\alpha, \nu) = \frac{\boldsymbol{E}\{(e_{\boldsymbol{R}}(\alpha) + \mu_e)(e_{\boldsymbol{R}}(\alpha + \nu) - \mu_e)\}}{\sigma_e^2}$$

Find the set $\{p_i^e : i = 1, \ldots, n_b\}$ most prominent peaks of $\tilde{R}_{\boldsymbol{e}}(\alpha)$.
if $n_b = 1$ **then**
  Set $h = 0$.
else if $n_b = 2$ **then**
  if $p_2 < \alpha_f - 2\lambda_\delta$ **then**
    Set $h = 0$;
  else if $p_2 < 2\alpha_f - \lambda_\delta$ **then**
    Second peak is most likely to relate the second harmonic, set $h = 2$.
  else if $p_2 < 3\alpha_f - \lambda_\delta$ **then**
    Second peak is most likely to relate the second harmonic, set $h = 3$.
    $\vdots$
  else
    Set $h = \left\lfloor \mu_\delta / \alpha_f \right\rfloor$.
  end if
else if $n_b > 2$ **then**
  $k = 1$
  while $i <= n_b$ **do**
    if $p_i < \alpha_f - i\lambda_\delta$ **then**
      Set $h = 0$; Break while.
    else if $p_i < (i+1)\alpha_f - \lambda_\delta$ **then**
      Set $h = 2$. Break while.
    else if $p_i < (i+2)\alpha_f - \lambda_\delta$ **then**
      Set $h = 3$. Break while.
      $\vdots$
    else
      Set $h = \left\lfloor \mu_\delta / \alpha_f \right\rfloor$.
      if $h < \alpha_f - \lambda_\delta$ **then**
        Set $h = 0$. Break while.
      end if
      $k = k + 1$.
    end if
  end while
end if
if $h = 0$ **then**
  No identifiable harmonic was found due to estimation problems of the STCFS.
else
  $\delta(t) = \delta_{\boldsymbol{y}}^0(t) / h$.
end if
Return $\delta_{\boldsymbol{y}}(t)$.

noise detection arises as a task to identify those signals that are contaminated with components that are not related with the cyclo–stationary process in order to detect time intervals over which the signal is too corrupted to provide information, or signals that have too low signal to noise ratio and any other processing task could not be reliable given this situation. Mainly there are two noise or interferences sources that disrupt the natural periodicity of the non–stationary process: *i) High frequency interference* and *ii) in band time-frequency interferences*.

4.6.1 High Frequency Interference Detection

Cyclo–stationary process usually have a defined frequency spectrum, i.e. phonocardiographic signals fundamental frequencies are to be below the $150Hz$ but are even more specific according to the heart age, rotating machines cycle frequencies are defined by the load and discs size, etc. Thus, the fact that external factors modify the cycle frequencies poses the necessity for a framework methodology that adapts the cyclo–stationary analysis to this external constraints. In [24], an interference detection for phonocardiographic (PCG) signals is proposed using fixed length windows to perform the detection of high frequency interferences over the signal.

Given a frequency range of interest $[f_l, f_h]$ for a given cyclo–stationary process $\boldsymbol{y} \in \mathbb{R}^T$, any short duration bursting components whose frequency range exceeds the maximum frequency allowed or assumed to hold process information is considered high frequency interference. In [23], this high frequency interference is detected by computing the energy of short time portions of signal and comparing this value either against the total energy signal or against the one-cycle energy, this is,

$$e^i_{\boldsymbol{y}} - \tilde{e}_{\boldsymbol{y}} \geq \eta_h \tag{4-18}$$

where $e^i_{\boldsymbol{y}}$ is the signal energy at a fixed length segment i and $\tilde{e}_{\boldsymbol{y}}$ can be either the average signal energy per cycle (note that a necessary condition here is to know the number of cycles in the analyzed signal duration) or the total signal energy, that is computed as the Shannon's energy as follows [11],

$$e^i = \sum_{\forall k \in [a_i, b_i]} -y\,[k]^2 \log\left(y\,[k]^2\right) \tag{4-19a}$$

$$\tilde{e}_{\boldsymbol{y}} = \boldsymbol{E}\{e^i_{\boldsymbol{y}} : \forall i \in 1, \ldots, n_h\} \tag{4-19b}$$

where $[a, b]$ is a non-overlapping signal interval, such that the segment $i + 1$ is contained

within $a_i = b_{i-1} + 1$ and $b_i = b_{i-1} + T_l$ with T_l the minimum window size for detecting bursting energy peaks higher than the maximum frequency allowed for the cyclo–stationary process and $n_h = \lfloor T/(T_l + 1) \rfloor$ and $\lfloor \cdot \rfloor$ is the nearest integer approximate of its argument.

4.6.2 Time-Frequency Analysis for In-band Interference Detection

Additional to the aforementioned high frequency interferences, cyclo–stationarity is usually affected by noise and artifacts that overlap with the natural frequency spectrum of the periodic non–stationary process. This fact makes difficult and sometimes unfeasible the extraction and filtering of the information related components. Then an identification procedure that takes into consideration the cyclic nature of the evolutionary process to detect the existence of corrupting components is presented in [24]. The process takes advantage of an enhanced subspace representation of the data to detect behavioral changes of the autocorrelation function over different non–overlapping frequency bands.
Let $y(t)$ be a cyclo–stationary process with fundamental cycle frequency α_i, such that in a time window T_t, that covers at least 3 complete cycles is possible. And let $\Omega(t, \omega)$ be its time-frequency representation constructed is such way that no cross-terms exist, two different time-spectral signatures are then posed as necessary condition for interference free signals:

- $\boldsymbol{E}\{\rho_{\boldsymbol{y}}(i, i+1)\} \leq \eta_\rho$ with $i = 1, \ldots, n_t - 1$
- $sign\left(\theta_{\boldsymbol{y}^i}^{(j)} - \theta_{\boldsymbol{y}^i}^{(j+1)}\right) = \text{const}$ with $i = 1, \ldots, n_t$ and $j = 1, \ldots, n_f$.

where $n_t = \lfloor T/(T_t) \rfloor$ and n_f the number of desired frequency ranges.
To this end, let the energy envelope of the given cyclo–stationary process be,

$$\epsilon_{\boldsymbol{y}}(t) = \sum_{\forall \omega} |\Omega(t, \omega)|^2 \tag{4-20}$$

The most prominent peaks of the energy envelope identify those components related to the fundamental components of the cyclic behavior, such that the similarity among cycles is used as an indicator of the existence of corrupting components. To this end, the radial measure $\rho_{\boldsymbol{y}}(\boldsymbol{y}^i, \boldsymbol{y}^{i+1})$ among the energy autocorrelation function of two consecutive non–overlapping signal segments that contains at least 2 complete cycles is used,

$$\rho_{\boldsymbol{y}}\left(\boldsymbol{y}^i, \boldsymbol{y}^{i+1}\right) = \frac{\langle R_{\boldsymbol{y}^i}(t^i, \tau), R_{\boldsymbol{y}^{i+1}}(t^{i+1}, \tau) \rangle}{\left\| R_{\boldsymbol{y}^i}(t^i, \tau) \right\| \left\| R_{\boldsymbol{y}^{i+1}}(t^{i+1}, \tau) \right\|} \tag{4-21}$$

where $i = 1, \ldots, n_t$ is the number of available consecutive non–overlapping windows in T. Hence, whenever the value of the radial distance exceeds a given threshold η_ρ. It should

be notice that the window size must be fixed prior to any procedure and thus requires of previous knowledge or estimation of the fundamental cycle frequency.

In addition to the temporal signature that measures how similar or dissimilar are consecutive segments or cycles of the processed signal, spectral similarity also serves as an indicator of interference components that disrupt the normal cyclo–stationarity of the process at hand in a given segment. Spectral similarity is then quantified by means of the peak alignment over different non-overlapping consecutive spectral bands distributed along the frequency spectrum of the cyclo–stationary signal.

That is, for a frequency range bounded by the process features with low and high frequencies $[f_l, f_h]$, we define a fixed number of bands of interest $1, \ldots, n_f$, then the enhanced time-frequency representation $\Omega_{\boldsymbol{y}}(t, f)$ over the frequency range of the given signal $y(t)$ is divided into n_f different sub-bands representations, for each of which the autocorrelation function of the energy envelope is obtained as in equation (4-20) and stacked all together into a matrix form,

$$\boldsymbol{A}^i = \left[R^1_{\boldsymbol{y}^i}(t,\tau), R^2_{\boldsymbol{y}^i}(t,\tau), \cdots, R^{n_f}_{\boldsymbol{y}^i}\right]^{\mathsf{T}} \tag{4-22}$$

where the superindex $i = 1, \ldots, n_t$. In [24] its pointed out that for completely cyclo–stationary processes the matrix $\boldsymbol{A}$ should demonstrate a strong linear dependency among rows, since peaks of each band envelope autocorrelation function must be almost perfectly aligned one to another and the peaks shapes within bands should exhibit a monotonic behavior that can be exemplified by means of the eigenvalues of the matrix $\boldsymbol{A}$. However, since real processes are not completely but instead almost cyclo–stationary, this condition is relaxed by computing the eigenvalues of sub-matrixes of matrix $\boldsymbol{A}$, that is, linear dependency and monotonicity is checked within fixed portions of the signal spectrum instead of the full spectrum at once:

$$\boldsymbol{A}^i_j = \left[R^j_{\boldsymbol{y}^i}, R^{j+1}_{\boldsymbol{y}^i}, \cdots, R^{j+n_b}_{\boldsymbol{y}^i}\right]^{\mathsf{T}} \tag{4-23}$$

with $j = 1, (n_f - kn_b) + 1, \ldots, n_f - n_b + 1$ and $k = \lfloor n_f/n_b \rfloor$.

Thus, by means of the eigenvalue decomposition set $\left\{\lambda^{(j)}_k\right\} \forall k = 1, \ldots, n_b$ of each sub-matrix, where k indicates the eigenvalue place in descending order associated with the j sub-matrix using the time-spectral signature is obtained:

$$\theta^i_\lambda(j) = \left(\frac{\lambda^{(j)}_2}{\lambda^{(j)}_1}\right)^2 \tag{4-24}$$

The time-frequency signature based on cyclo–stationarity features requires then that the spectral signatures $\rho^i_\lambda(j)$ holds a given monotonicity either by increasing or decreasing values values as the spectral frequency bands increases, so that one of the following relationships must hold,

$$\theta^i_\lambda(j) < \theta^i_\lambda(j+1) < \ldots < \theta^i_\lambda(n_f) \tag{4-25a}$$

$$\theta^i_\lambda(j) > \theta^i_\lambda(j+1) > \ldots > \theta^i_\lambda(n_f) \tag{4-25b}$$

If neither of the previous conditions are accomplished, the signal segment is assumed to hold spectral non–stationarities that do not meet cyclo–stationary constraints and therefore signal is corrupted, that is, if neither (4-25a) or (4-25b) holds, the parameter criteria is set $\eta_\theta = 1$, on the other hand $\eta_\theta = 0$.

4.7 An Enhanced High Frequency Interference Detection based of Adaptive Window size

Usually, high frequency interference detection is done prior any other analysis by means of sliding non–overlapping windows and the Shannon's energy of equation (4-19a). Yet, we propose using a adaptive window size computed using external information and a bias sigmoid curve that adjust the size value so that existence of complete cycles is fulfilled. Thereby equations (4-19a) and (4-19b) become,

$$n_h(a) = \left\lfloor \frac{T}{T_{\alpha_h} + 1} \right\rfloor \tag{4-26a}$$

$$T_{\alpha_h}(a, \beta_a, \kappa_a, c_a) = \beta_a + \frac{1 - \beta_a}{1 + e^{-\kappa_a(a - c_a)}} \tag{4-26b}$$

where the parameter a is fixed according to external factors that regulate the maximum expected frequencies, β_a sets the lower asymptote value, κ_a sets the value for which the function starts growing and c_a the settling value for the higher asymptote. An example of a sigmoid function to regulate window size in phonocardiographic signals based on age is shown in figure **4-1**,

After computing signal segments energy, same condition as in static procedure is applied, such that any segment that exceeds the predefined η_h threshold value is marked as a bursting energy component that does not relate with the cyclo–stationary process and whenever the number of segments is considered too large signal is considered as too corrupted for reliable

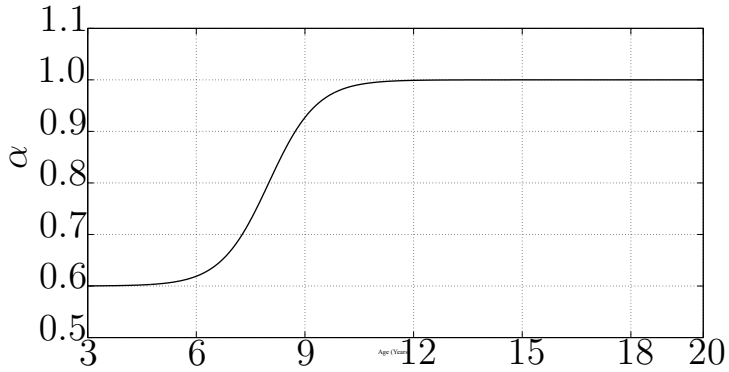


Figure 4-1: Sigmoid Penalty Function for PCG signals using age a as penalty factor, $\beta_a = 0.6$, $\kappa_a = 1.5$ and $c_a = 8$

analysis.

4.8 Time-Frequency Analysis for In-Band Interference Detection using Adaptive Constraints

Almost cyclo–stationary processes do not hold an static and specific cycle frequency over time, instead the cycle frequency can present a drift over the real estimated value do to external factors that have not been taken into consideration so far, i.e. in cardiological related signals, the Heart Rate Variability (HRV) influences the length of the cardiac cycle trough time causing non–periodic increase or decrease in the cycle frequency while for rotating monitoring systems and vibrations signals, such drift might be caused by an increase in the number of revolutions per minute, a fault in the mechanisms among others. Additionally to this drawback, the necessary condition of the number of cycles within the signal generates either supervised monitoring process, or the necessity of additional information, such as exact cycle frequency with minimum deviation or additional simultaneously recorded signals that provide such information (such as electrocardiogram for PCG signals). To overcome this, we propose an adaptive interference detection scheme that uses non–supervised estimation of cycle frequencies and window size estimation based upon external factors and cyclic features. As in section §4.6.2, consecutive cycles are compared by means of a similarity measure computed as the radial distance among the energy autocorrelation functions. However the necessary condition of at least 3 complete cycles within the segment of analysis must not be fulfilled given the time-variant nature of the cycle frequency, not mentioning that to any change of signal process such window must be reseated to this constraint, or just work under the wrongful assumption of this number of cycles. Similar to the adaptive high frequency interference detection an adaptive window size parameter is here introduced as an alternative

to this drawback, not only using external information that can be automatically set but also taking into consideration the time-variant behavior of the cycle frequency obtained through algorithm 1.

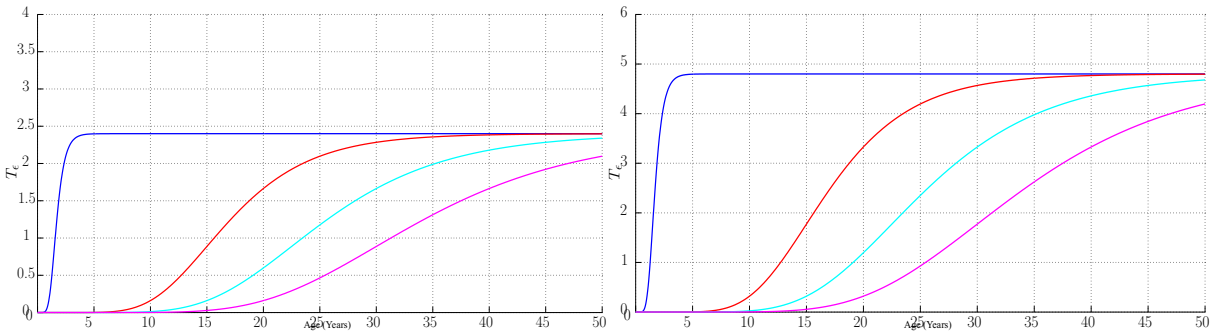
For a cyclo–stationary process $y(t)$ which holds time-variant cycle frequency $\alpha_i(t)$, the Gompertz function [28] is used as adaptive time signature window,

$$T_t(a_t, \kappa_t, \beta_t, c_t) = \kappa_t e^{(\beta_t e^{c_t a_t})} \tag{4-27}$$

where κ_t sets the upper asymptote value, $\beta_t < 0$ sets the x-axis displacement, $c_t < 0$ sets the growing rate and a_t is a external related parameter fixed accordingly to the maximum expected frequency of the cyclo–stationary process. To incorporate the variant information of the cycle frequency, parameters κ_t and c_t are interchanged by four times the value of the cycle frequency at sample point of analysis and its standard deviation among the whole available duration of the process respectively.

It should be remarked that instead of the common approach which requires a prior fixed length for any window based analysis, the adaptive windows can evolve with the process nature, such as if the external parameters change the temporal computation of the windows size is capable to detect and adapt itself to this changes.

An example of this function with different combinations of the parameters is shown in Figure **4-2** for PCG signals, where the parameter a_t is the age of the heart muscle.



(a) T_ϵ for $\mu_{\tau_d} = 0.6$. (b) T_ϵ for $\mu_{\tau_d} = 0.9$.

Figure 4-2: Sigmoid Functions for Adaptive Time-Frequency Signatures Windows Estimation■– $\sigma_{\tau_d} = 0.005$, ■– $\sigma_{\tau_d} = 0.05$, ■– $\sigma_{\tau_d} = 0.075$ and ■– $\sigma_{\tau_d} = 0.1$

4.8.1 Peak Alignment and Spectral Autocorrelation Coherence

Peak alignment is also considered as a criteria for interference presence, however, taking the first frequency band autocorrelation function, proposed by [24], as reference might lead

to false peaks if there is major interference in this band, thereby a proper selection of the reference band is required, we propose a correlation based approach, with the average correlation coefficient criteria, $\varrho_i = \frac{1}{n_b}\sum_{j=1}^{n_b} R_{ij}$, where $\boldsymbol{R}$ is the correlation coefficients matrix:

$$R_{ij}\left(e_x^i\left(T_t\right), e_x^j\left(T_t\right)\right) = \frac{\boldsymbol{E}\{e_x^i\left(T_t\right)e_x^j\left(T_t\right)\} - \boldsymbol{E}\{e_x^i\left(T_t\right)\}\boldsymbol{E}\{e_x^j\left(T_t\right)\}}{\left(\boldsymbol{E}\{\left(e_x^i\left(T_t\right)\right)^2\} - \boldsymbol{E}\{e_x^i\left(T_t\right)\}^2\right)^{\frac{1}{2}}\left(\boldsymbol{E}\{\left(e_x^j\left(T_t\right)\right)^2\} - \boldsymbol{E}\{e_x^j\left(T_t\right)\}^2\right)^{\frac{1}{2}}}$$

The reference band index is then $I_r = \arg\max_i\{\varrho_i\}$. Having defined the reference band autocorrelation function, the most prominent peaks of each band are extracted and compared against those of the reference, a peak of the autocorrelation function r_x^j is considered aligned if its within $75ms$ of any peak in the reference autocorrelation function $r_x^{I_r}$, i.e., let $\boldsymbol{p}^j \in \mathbb{R}^{1\times M}$ and $\boldsymbol{p}^{I_r}\ in\,\mathbb{R}^{1\times N}$ be the most prominent peaks of the $j-$th band and the reference band respectively, the peak alignment criteria is defined as $\psi_p = \frac{1}{n_b}\sum_{j=1}^{n_b}\boldsymbol{E}\{\boldsymbol{p}^j\}$, where,

$$p_m^j = \begin{cases} 1 & \left\|p_m^j - p_n^{I_r}\right\|_2 \leq \eta_\psi \quad \forall n = 1,\ldots,N \\ 0 & \text{Otherwise} \end{cases} \tag{4-28}$$

Finally, in addition to temporal and spectral periodicity criteria, the spectral and temporal energies of the signal segments are also indicators of not only non-cyclic related events, but also of signal amplitude attenuations. Therefore, the correlation coefficient between the time-series segments and a reference cycle segment (chosen within the signal with duration T_t, periodicity parameter $\rho_\chi^{I_r} > 0.8$, $\rho_\theta^{I_r} = 1$ and percentage of align peaks $\geq 90\%$), is used in a template matching procedure, where the correlation coefficient $\chi\left(\Omega_{rms}^{ref}\left(f\right), \Omega_{rms}^{j}\left(f\right)\right)$ among the frequency marginal of the reference segment and the frequency marginal of the analyzed segment is considered as noise free if $\chi \geq \eta_\chi$, where η_χ is a threshold value, and the frequency marginal is defined as,

$$\Omega_{rms}^i\left(f\right) = \sqrt{\int \left|\Omega^i\left(f,t\right)\right|^2 dt} \tag{4-29}$$

where $S^i\left(f,t\right)$ is a time-frequency representation of the segment $y^i\left(t\right)$.

5. Validation Results

In this section, real data time-series are used in order to assess and validate the cyclo–stationarity analysis through the estimation of time–variant cycle frequency estimation under different noise conditions. To this end, phonocardiographic signal (PCG), an acoustic cyclo–stationary process related to the mechanical function of the heart muscle, is used.
PCG time–series cyclo–stationary analysis and interference detection based on cyclo–stationary constraints is aimed to enhance segmentation procedures required as pre-processing stage, given the wide signal spectrum such interference can not be filtered from the process without loss of information [8], thus interference detection process would result a signal quality assessment methodology capable of identifying interference–free signals. Performance is evaluated in terms of interference detection capability under different noise conditions, accuracy of the estimated cycle frequency and enhancement in the segmentation procedure comparing between the conventional non-constrained framework and the proposed adaptive framework.

5.1 Non–Stationarity Analysis in PCG Time-Series

5.1.1 Database

The used heart sound database for murmur detection belongs to the Signal Processing and Recognition Research Group at the Universidad Nacional de Colombia-Manizales. The database holds records of 60 patients: 27 patients ages 3 to 10, 16 patients ages 10 to 14 and 17 patients 14 years old and onwards.
From each patient a 20 second duration phonocardiographic (PCG) signal was recorded in each of the four traditional auscultation focuses (aortic, mitral, pulmonary and tricuspid) using both, bell and diaphragm, modes. Some signals were dismissed due to acquisition quality but holding at least one signal per focus[1]. Signals were recorded at a $44.1kHz$ sampling rate and 16 bit resolution with an electronic stethoscope and Meditron® software, the acquisition was done in noisy environments as those of health care centers. The database was then divided into 4 sets according to each age range each holding 4 subsets according to

[1] A total of 453 signals from the 453 are available for segmentation evaluation performance

the auscultation focuses; each signal hold the annotations marks of the onsets and endings of the hearts sounds present in each signal so an estimated of the heart rate and by inference the average cycle frequency, Table 5-1 summarizes the database setting.

Table 5-1: Database Description

Ages	**Focuses**				
	Aortic	**Mitral**	**Pulmonary**	**Tricuspide**	**Total**
0 to 3	16	17	14	11	58
3 to 10	51	48	45	38	182
10 to 14	25	25	25	25	100
14 and older	30	27	30	26	113
Total	122	117	114	100	453

PCG signal present components in a lower frequency range than the original sampling frequency, therefore a 8-th order Tchebychev filter with a down-sampling process is implemented with a central cut-off frequency $f_c = 1102.5$, that is a down-sampling factor of 20. Thus resulting in a sampling frequency $\hat{f}_s = 2205Hz$. Finally the amplitude of each signal was normalized with the absolute maximum of each PCG recording.

5.1.2 Performance Measures

The performance of the proposed cyclo–stationary analysis for detecting the presence of interferences is assessed in two stages: first, by means of sensitivity the cycle–frequency estimation:

$$s_e = \frac{i_c}{i_c + i_u} \tag{5-1}$$

where i_c and i_u are the correct and uncorrect number of ICF estimations.
Secondly as the database does not hold any kind of annotations regarding the quality of the signals, a first process of database depuration is performed using hard interference detection parameters with the proposed scheme (which is depicted in Figure 5-1).
This provides with a reduced database with the cleanest signals available. Such signals are used to test the robustness of the methodology under different signal-to-noise ratio levels to detect the presence of interferences, to this, four types of additive noise processes are considered:

- Chirp signal: frequency modulated cosine varying frequency from $500Hz$ up to $900Hz$, to simulate bursting high frequency components.

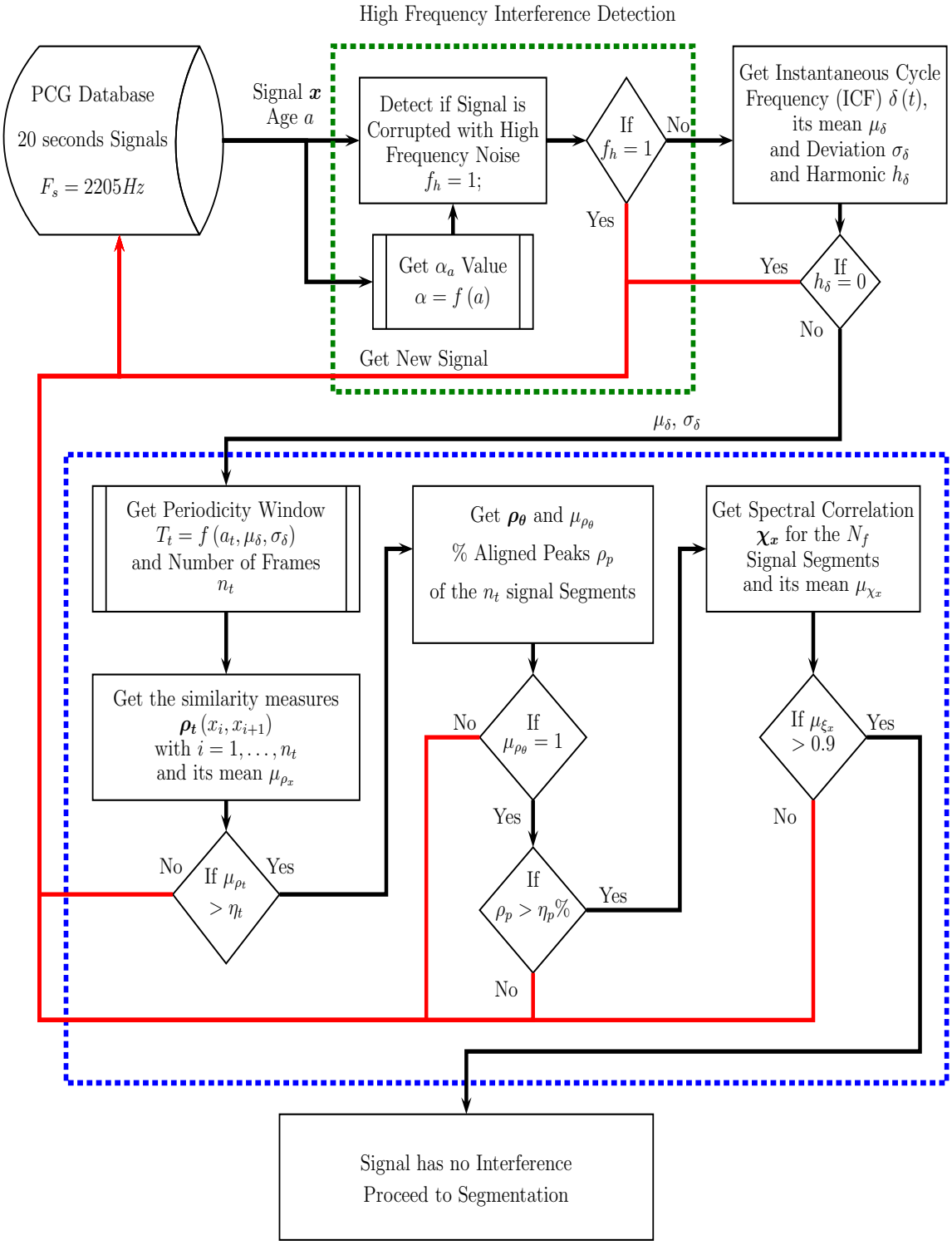


Figure 5-1: PCG Signal Noise Analysis Scheme.

– Colored gaussian noise to overlap PCG natural frequency spectrum, band $25-400Hz$.

– Voice signals of a moving loud speaker, recorded with two microphones, [4].

The noisy signal is constructed by linearly adding portions of noise to randomly selected segments to the original PCG signal. The sliding nature of the analysis provides then with annotations of corrupted signal segments, these are used to compute analyzed the detection capability of the conventional and adaptive frameworks.
Finally, the segmentation procedure of algorithm 2 presented in [6], based on temporal constraints is used to assess the quality of the detection[2].

Algorithm 2 Segmentation Procedure

Input PCG signal $x(t)$, ICF $\delta(t)$, $t_o = 0, t_f = 3$, set $i = 1$.
Get real onset annotations ξ.
while $t_f < 20$ **do**
  $x_i = x(\tau)$, $\tau = \{t_o, t_f\}$.
  Get Cycle duration at t_o, $C_i = 1/\delta(t_o)$ and $\epsilon_{x(\tau)}$ and find its prominent peaks.
  Check synchronism conditions.
  if The conditions hold **then**
    $\widehat{\xi}_i = p_1 - 0.15\mu_{\delta(t)}$
  else
    $\widehat{\xi}_i = p_2 - 0.075\mu_{\delta(t)}$
  end if
  $t_o = \widehat{\xi}_i + C_i$, $t_f = t_o + 3$.
end while
Compute the true and false detection probability using $\boldsymbol{\xi}$ and $\widehat{\boldsymbol{\xi}}$.

Segmentation performance is assessed by means of conventional true and false detection probabilities [32],

$$P_d = \frac{N_D}{N_D + N_M}, \qquad P_f = \frac{N_F}{N_D + N_F} \tag{5-2}$$

where N_D are the correctly detected points, N_F the number of false points, i.e. a point $\widehat{\zeta}$ that is marked as onset but has no correspondence with any of the true points $\boldsymbol{\zeta}$, and N_M a missed point.
As the true segments are manually annotated, a threshold of 15% of the heart rate is used when comparing the estimated onset and endings of PCG heart sounds to thus of the reference annotations. Figure **5-2** shows the process of evaluation for segmentation.

[2]Given a noise-free signal is segmented, taking segments lasting $3s$ to calculate a time-frequency subspace representation and its respective energy envelope $\epsilon(t)$. The TFR is chosen to be a continuous wavelet

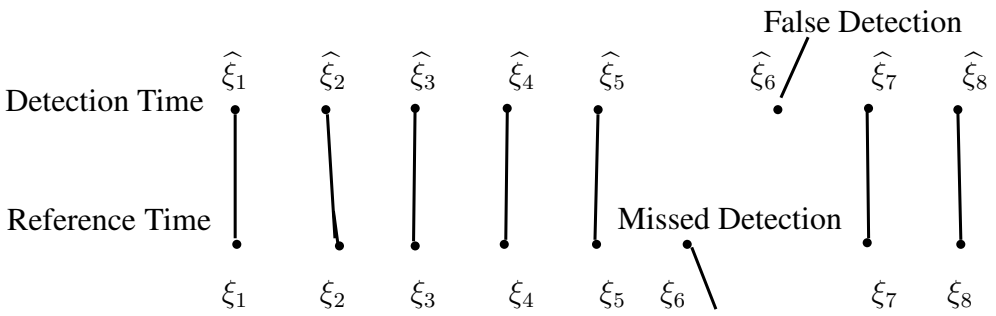


Figure 5-2: Segmentation evaluation procedure.

5.1.3 Cycle Frequency Estimation of PCG signals

Cycle frequency of cardiac related events is estimated as the inverse of the cardiac cycle duration, so that the number of complete cardiac cycles counted as the number of both onset and ending of the hearts sounds is used divided by the duration of the signal, provides with a general average cycle frequency. However, the cardiac cycle length might vary under several physiological conditions, usually known as heart rate variability (HRV) thereby, this average estimate cannot be directly compared with the average of the instantaneous cycle frequency. Hence, a permission threshold of 15% is used when analyzing the performance measure of the cycle frequency estimation of PCG signals. Additionally, the α_{max} value is manually set to $5Hz$, given that the maximum cycle frequency for PCG signals is expected to be less than $2.5Hz$ but since signals are expected to be corrupted by noise, the estimated ICF might be more clear over some harmonic value of the real ICF. The search deviation for ICF algorithm 1 is set $\lambda_\delta = 0.5Hz$, allowing a high rate variation within the 20s signal duration[36].

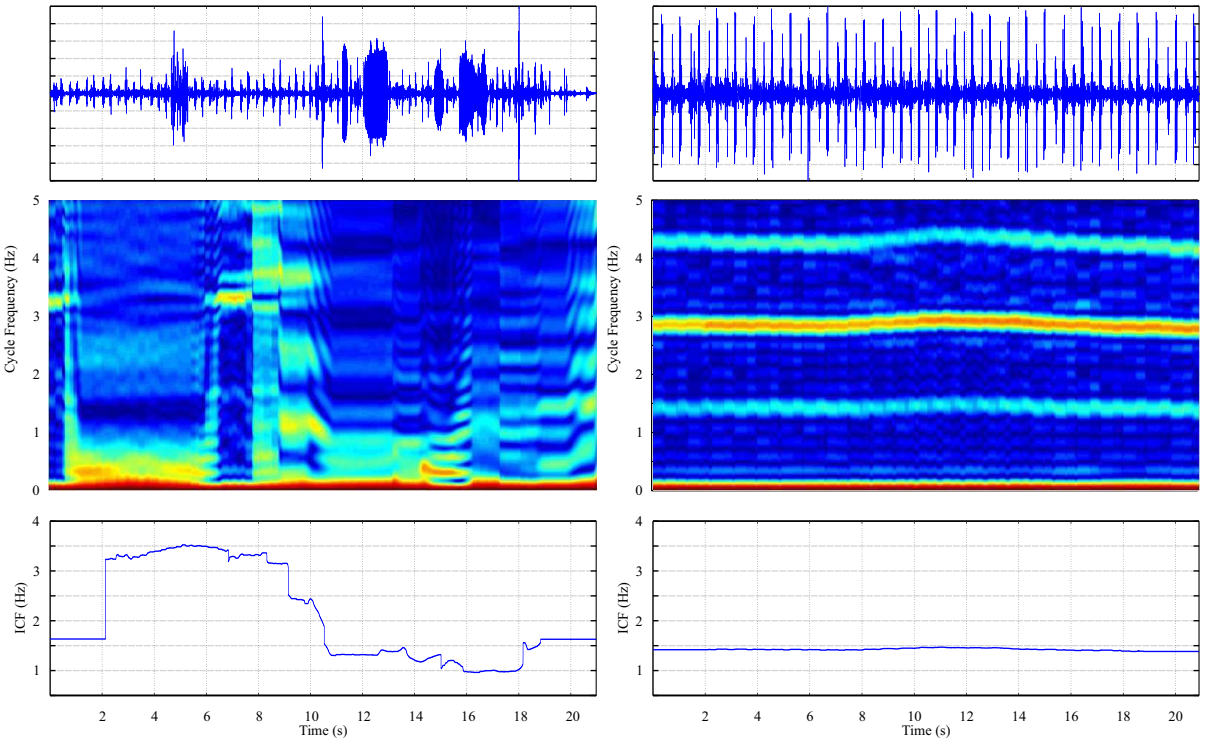
The proposed approach that uses the adaptive window length under age constraint (method **A**) is compared against the regular fixed size window (method **F**) with a length of 2s. Results are shown in Table **5-2**, where it can be seen a slight improvement in the sensitivity values for all PCG signals.

It must be remarked that major improvement is detected in auscultatory focuses with more interference presence, such as pulmonary and tricuspid, where respiratory and gastro-intestinal noises usually disrupt the normal heart sound. Two examples of noise contaminated signal and clean signal are shown in Figure **5-3**, first row depicts the PCG signals, second row shows the STCFS and last row shown the estimated ICF, it is clearly visible that interference existence disrupt the cycle time-frequency map and thus no cycle frequency estimation is possible.

transform (CWT) computed with a complex Morlet wavelet, and scales a are chosen so the frequency range covers the $25 - 600Hz$ bandwidth.

Table 5-2: ICF Average Sensitivity (s_e) per Focus

Ages	Focuses							
	Aortic		Mitral		Pulmonary		Tricuspid	
	F	**A**	**F**	**A**	**F**	**A**	**F**	**A**
0 to 3	0.94	**0.94**	0.88	**0.94**	0.79	**0.93**	0.91	**0.91**
3 to 10	0.90	**0.94**	0.92	**0.94**	0.91	**1.00**	0.87	**0.95**
10 to 14	**0.88**	0.84	0.88	**0.88**	0.88	**0.92**	0.84	**0.84**
14 and older	0.90	**0.90**	0.81	**0.93**	0.90	**0.97**	0.85	**0.92**



(a) Noisy Signal.

(b) Clean Signal.

Figure 5-3: STCFS and ICF estimation for contaminated and clean signals.

5.1.4 Interference Detection in PCG Signals

Noise Detection from high frequency components

The analysis window for high frequency interference detection in the conventional non–constrained approach is manually set to $100ms$, while the proposed approach uses the adaptive window length to constraint this window using equation 4-26a. The threshold parameter is set to $\eta_h = 0.075$, thus if in either of the two methods, the energy of a given window exceeds such threshold, it is marked as noisy, whereas if more than 10 segments are marked as noisy, the signal is marked as low quality due to high frequency components and thereby rejected.

Table 5-3: High Frequency Interference Detection Results.

Ages	Focuses							
	Aortic		Mitral		Pulmonary		Tricuspide	
	F	A	F	A	F	A	F	A
0 to 3	4	5	3	4	5	8	3	5
3 to 10	6	10	7	10	8	12	1	2
10 to 14	1	1	0	0	2	2	1	0
14 and older	1	1	6	6	4	4	2	2

Table 5-3 shows the rejected signals per age per focus using both methodologies, it can be seen, that adaptive framework rejects a higher number of signals (72) compared against conventional 54, specifically is the lower age ranges, when observing such signals its was appreciated a large number of speech components such as crying and voice. An example of such detection is shown in figure 5-4, where the red depicted segments are marked as noisy due to a voice detected component.

Even when both methodologies are capable of detecting the segment as noisy, static analysis does alert the existence of noise and assessed the signal as clean while the adaptive window analysis adjust itself to determine that this component does not meet energy criteria and marked the signal as noisy.

In-band Interference Detection

In-band interference detection is carried out over all signals no matter if the high frequency analysis had detected them as noisy, the only condition to perform this stage was a correct assessment of the cycle frequency, such that the window lengths used for analysis gave a proper value (3 full cardiac cycles required). This stage was performed using the parameters

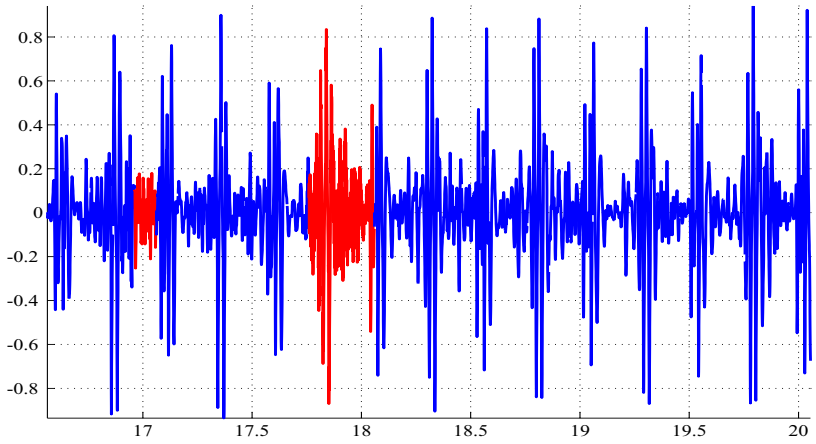


Figure 5-4: PCG Segment with High Frequency Noise.

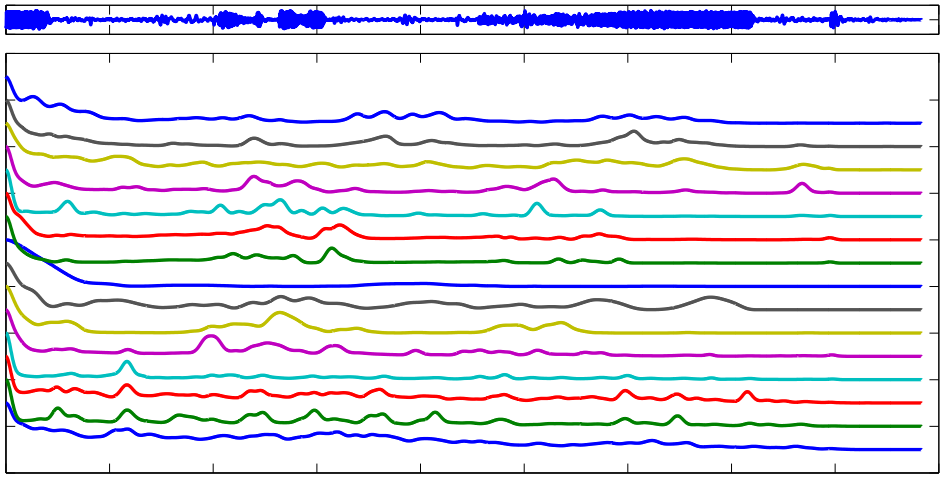
stated in Table **5-4**

Table 5-4: In-band Analysis Parameterization

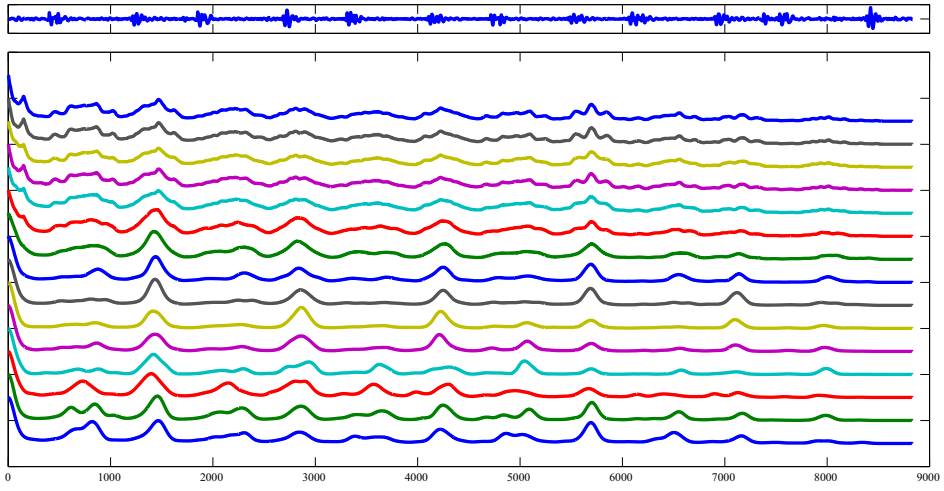
Signature	**Window**		**Threshold**	
	F	**A**	**F**	**A**
Temporal	$T_t = 2.5\text{s}$ $n_f = 15$ $n_b = 3$	$T_t\left(a, 2.4, 4\mu_{\delta(t)}, -10, -0.01\sigma_{\delta(t)^{-1}}, \alpha_a\,(0.6)\right)$ $n_f = 25$ $n_b = 5$	$\eta_\rho = 0.7$	
Spectral			$\eta_\theta = 0.8$	
Alignment			$\eta_\psi = 0.9$	
Coherence			$\eta_\chi = 0.8$	

Figure **5-5** shows an example of noise 5.5(a) and clean 5.5(b) PCG signal segments, it is visible how interferences affect the homogeneity of the energy autocorrelation function over the frequency bands, on the other hand, for clean segments, the energy autocorrelation is very similar among bands and peaks related to heart sound are quite well aligned. There is to notice in figure 5.5(b), that the first frequency band does not hold the peak alignment reference, since there is not enough energy concentration over this band, whereas second or third band serve as a better template.

In–band detection analysis, even when threshold were fixed equally for conventional and constrained approach, resulted in a different number of rejected signals, while the first one rejected a lower number of signals, however segmentation performance also resulted in lower accuracy values. Table **5-5** shows the rejection results for in-band interference, same as in the high frequency case, there is a increment in the number of rejected signals in the lower age ranges, an additional reason for this, is that the cardiothoracic apparatus on infants is less develop and small, which leads to recording several internal sounds related to gastrointestinal function and pulmonary and respiratory interferences. Moreover, the higher



(a) Noisy Signal.



(b) Clean Signal.

Figure 5-5: Spectral Signatures for noisy and clean PCG segments.

Table 5-5: In–Band Rejected Signals

Ages	Focuses							
	Aortic		Mitral		Pulmonary		Tricuspide	
	F	A	F	A	F	A	F	A
0 to 3	12	13	9	10	11	11	10	9
3 to 10	37	43	38	38	39	42	27	29
10 to 14	22	21	17	19	21	20	14	16
14 and older	23	23	20	19	19	24	19	18

frequency rates on infants makes the interferences to disrupt the normal silences of the heart sounds, such that temporal similarity has low values and spectral coherence its not possible. The segmentation performance, done as explained in [6], results are shown in Table **5-6**

Table 5-6: Segmentation Performance

(a) True Detection Probability.

Ages	Focuses							
	Aortic		Mitral		Pulmonary		Tricuspide	
	F	A	F	A	F	A	F	A
0 to 3	0.69	**0.88**	**0.67**	0.60	0.70	**0.70**	0.85	**0.87**
3 to 10	0.62	**0.79**	0.62	**0.61**	0.74	**0.74**	0.64	**0.69**
10 to 14	0.67	**0.74**	0.58	**0.76**	**0.79**	0.76	0.64	**0.73**
14 and older	0.77	**0.78**	0.51	**0.58**	0.61	**0.64**	0.72	**0.76**

(b) False Detection Probability.

Ages	Focuses							
	Aortic		Mitral		Pulmonary		Tricuspide	
	F	A	F	A	F	A	F	A
0 to 3	0.19	**0.04**	**0.22**	0.30	0.12	**0.12**	0.00	**0.02**
3 to 10	0.23	**0.07**	**0.17**	0.25	**0.06**	0.15	**0.12**	0.14
10 to 14	0.15	**0.09**	0.21	**0.04**	**0.01**	0.22	0.19	**0.05**
14 and older	**0.03**	0.13	0.32	**0.25**	0.15	**0.25**	0.13	**0.06**

Regarding segmentation performance, even though the number of signals of adaptive non–stationary PCG analysis is lower, it is important to remark that there is a general increase in the average true detection probabilities, both when doing a age ranked analysis or an auscultation focus. Not mentioning that segmentation procedure its only based on temporal conditions and durations of heart sounds, which indicates that a proper quality assessment prior to any task might improve analysis results.

5.1.5 Robustness of Cyclo–Stationary Signal Analysis and Non–Stationarities Detection

Finally, the last experiment is aimed to test the robustness of the interference detection methodologies, thus only the 98 non–rejected signals provided by the adaptive non–stationary analysis are used. Under the assumption that such signals have high signal-to-noise ratio (SNR), signals are contaminated with several noise sources in random segments using four different SNR values $0, 3, 6, 9 dB$. The methodology is tested for all stages in a sequential procedure, first high frequency interference, followed by the ICF estimation and lastly the

in-band analysis. Table 5-7 shows the rejected signals for conventional (top) and proposed (bottom) analysis, numbers one to 4 on the second row of both tables identify the auscultation focuses: **1: Aortic**, **2: Mitral**, **3: Pulmonary** and **4: Tricuspide**.

Table 5-7: Noise Analysis

(a) Conventional Analysis, fixed size window.

Ages	**0dB**				**3dB**				**6dB**				**9dB**			
	1	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**
0 to 3	3	7	3	2	3	7	3	2	3	7	3	2	2	7	3	1
3 to 10	8	10	3	9	8	10	3	9	8	10	3	9	8	9	3	7
10 to 14	4	6	5	9	4	6	5	9	4	6	5	9	4	6	5	9
14 and older	7	8	6	8	7	8	6	8	7	8	6	8	5	7	4	8

(b) Proposed Approach, adaptive window size.

Ages	**0dB**				**3dB**				**6dB**				**9dB**			
	1	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**
0 to 3	3	7	3	2	3	7	3	2	3	6	2	2	2	6	3	2
3 to 10	8	10	3	9	8	10	3	9	8	9	3	8	8	8	2	7
10 to 14	4	6	5	9	4	6	5	9	4	6	5	9	4	5	5	7
14 and older	7	8	6	8	7	8	6	8	6	6	6	8	5	5	2	5

The threshold parameters were equally set for both frameworks as: $\eta_h = 0.08, \eta_\rho = 0.7, n_f = 15, n_b = 3, \eta_\psi = 0.8\eta_\chi = 0.8$. This analysis its less permissive than the one done for the whole database given the high SNR assumption of the original signals. It can be seen from the table that, conventional fixed window analysis, its not capable of adjusting itself to the induced noise, since even for high values of SNR, it rejects most of the signals (88 out of 98 for $9dB$), that is, given that there is an existing interference, it is not capable of detect how much the interference affects the natural evolution of the time series and has no other option than to dismiss the signal. On the other hand, the adaptive analysis dismiss only 76 signals, so even when in those signal noise its detected, the thresholds criteria allows for the signal to be considered as clean, that is, noise is not that relevant for disrupting the temporal or spectral signatures.

Finally, the detection capability was also analyzed, observing that both frameworks were capable of detecting the 99% of the noise portions induced, however, since the adaptive windows, produced with more narrow and accurate windows, has a better localization ability than the fixed window analysis. This poses then, with early alert functions to the methodology, and might be useful in signal acquisition quality assessment.

6. Discussion and conclusion

This chapter presents a analysis framework for detecting the presence of non–stationary components in cyclo–stationary processes, the methodology poses several adaptive constraints that uses the internal cyclic behavior of the process to adjust commonly used parameters such as window sizes for the subspace representations (time-frequency representations, frequency sub-band analysis and energy envelope matching). Detection of non–stationarities is compared against non–adaptive methodology and improvement its shown in terms of visual analysis and posterior processing task. The obtained results, allows to conclude the following:

i) Introduced adaptive window parameterizations based on internal structure information must be remarked as an important tool when analyzing cyclo–stationary processes. To this end, several weighting function are used altogether with external information and internal structure characteristics so that, analysis takes into consideration the time-variant nature of the cyclic phenomena. Yet, it is important to have or extract some other information that relates not only the periodic evolution of the time–series, but also directly analyze the non–stationarities so that this components can be analyzed in some cases as part of the process information and are not directly dismissed as noise. Such is the case of other applications as vibration monitoring, where the fault diagnosis commonly uses blind component extraction technique for separating stationary, non–stationary and cyclo–stationary process, where the non–stationary process usually holds information of the fault sources as the signal mixtures is a convolutive process.

ii) Cyclo–stationary process analysis, using subspace based representations is carried out under a time–frequency approach. The main goal of behind the time-frequency representation is to enhance the spectral information related to the time–series, a proper selection of the TFR stuck out as primal stage, so that the proposed framework. Specifically, when analysing PCG signals, continuous wavelet transform is used, due to its high resolution for the specific frequency ranges, however, some other spectral representations might be useful, such as singular spectrum analysis or empirical mode decomposition. Additionally, the adaptive windows, can be posed into a online methodology that adapts to the time variance of cyclic features.

It is important to remark that sequential analysis of temporal and spectral signatures is not a necessary condition of the methodology, thresholds value can be adjusted to bias the results into one specific aspect, as an example, since segmentation procedures use temporal localization and duration of events, the time signature can be given higher threshold values, whereas if classification were the final application and spectral features were used, frequency signatures might have greater weight.

iii) State of the art methodology is compared with proposed constrained approach and evaluated in terms of posterior segmentation analysis for PCG signals. True and false detection probability values were improved when segmenting the clean signals assessed by adaptive methodology. To observe the detection capability, several noise sources were induced to some signals, results proved that adaptive analysis is capable of detecting noise for low and high values SNR values, yet in the last case, the sequential process allows to determine that interferences do not disrupt natural information of the cyclo–stationary process, while state of the art framework does not exhibit such property.

As conclusion, its has been presented a framework for adaptively analyzing the cyclo–stationary properties of time–series, assessing the presence of non–stationary components by means of time-frequency subspace representations. The method uses not only internal information but also external given parameters to adjust the involved parameters so that temporal and spectral analysis take into consideration time-variant characteristics. Besides, robustness analysis is tested with assumed high SNR value signals when added with several interferences and methodology proves to be efficient in assessing this presence under low signal quality but also determines that for higher, good enough SNR values, information its not compromised.

Future work includes then, using the framework for assessing and developing quality monitoring systems, so that automated online diagnosis and early alert of poor acquisition can be given. Additionally, accurate cycle frequency estimation based on adaptive windows is to be used in process of adaptive blind component extraction, where this value is usually manually fixed and does not take into consideration any time variant, which is a common property is digital signal processing task. Application of the proposed framework on vibration monitoring systems is also considered for online fault detection. Nonetheless such applications would require some other subspace based projection more suitable to such tasks, not mentioning that in some cases analysis of the stochastic properties from a time approach is not recommended and an purely spectral analysis adaption is required.

Part III

Detection and Extraction of Non-Stationary Components with subspace stationary constraints

7. Non–Stationary Component Extraction from Subspace Stationary analysis

Signal stationary analysis has been a milestone for digital signal processing and pattern recognition tasks. Filtering and extraction of linearly mixing models based on prior knowledge are of common requirement in multiple fields of analysis such as communications, biomedical signal analysis, econometrics and time-series data managements and neuroscience. In such context a major point of interest has been the signal separation of components that are assumed to be stationary under certain stochastic constraints. To this end, most common approaches uses weakly stationary conditions as the time invariability of second order statistics along time, i.e., constant mean and variance values through fixed periods of time for locally stationary process [5]. Yet, a major concern still relates to the blind extraction of the non–stationary signals in filtering separation tasks, where the parameter estimation lacks of physical interpretability. To overcome this drawback, some methods rely on the dynamic behavior of the stochastic internal structure for a given estimation model. In [7] an online Gaussian signal processing is used to simultaneously tackle the non–stationary and the temporal correlation of the mixture model. Whereas in [19] similar work is presented from a hidden Markov model processing. In [30] a different approach is used to separate non–stationary signals from a time-frequency based representation but under the necessary condition that distribution of mixed signals do not overlap in time and frequency.

In this chapter, the Analytic Stationary Subspace Analysis (ASSA) proposed in [20] is used and extended for one-dimensional time-series formed by linear mixtures of stationary and non–stationary processes aiming to develop a less restrictive separating approach based on conventional nonparametric subspace analysis. Stationary signal behavior is highlighted by means of enhanced data representations and stochastic constraints are put together in joint subspace analysis to result in projection spaces where signals are as stationary as possible. Results are shown using synthetic simulated time series where real and estimated stationary signals are compared by means for degree of correlation and biological real data, specifically

electroencephalographic signals, where a given measure of non-stationarity is used to assess the extracted components.

7.1 Signal Stationary Analysis using Subspace Based Methods

Recall equation (4-1) and similarly to cyclo–stationary process the formal definition of a weakly stationary process is given by

Definition 7.1.1. (*Weak Stationarity*) Let $y(t) \in \mathbb{R}^T$ be a random time–series, the stochastic model for weak stationarity of the given process is,

$$\lim_{\tau \to T} \boldsymbol{E}\{|\Theta\{y(t)\} - \Theta\{y(t+\tau)\} : \forall \tau \in T|\} - \epsilon \leq 0\ , \forall t \in T \tag{7-1}$$

where $\epsilon \subset \mathbb{R}^+$ is a given small value that relates the degree of stationary accuracy.

The main goal of conventional subspace analysis can defined as searching for a projection subspace $\boldsymbol{\Phi}$, bearing still as much information of the input process. The separation of stationary and non–stationary time–variant features can consequently be carried out by adding the stochastic constraint of Definition 7.1.1. For one dimensional time–series, i.e. $y(t) \in \mathbb{R}^T$, the blind component extraction of stationary and non–stationary signals yields a filtering–separation task defined as follows

Definition 7.1.2. *Signal Separability* Let $x_s(t) \subset \mathbb{R}$ denote a stationary time-series and $x_n(t) \subset \mathbb{R}$ be a corrupting non–stationary process. The measured observation process time series is assumed as the linear mixed process $x(t) = x_s(t) + x_n(t), t \in R \subset \mathbb{R}$. The signal separation problem is, by definition, to determine the conditions under which an estimate of the stationary process is obtained from the input measured process $x(t)$ to a given degree of accuracy, that is,

$$\mathcal{M}\{\boldsymbol{x}, \epsilon\} = \tilde{\boldsymbol{x}_s} + \tilde{\boldsymbol{x}_n} \tag{7-2}$$

where $\tilde{\boldsymbol{x}}_s$ is the estimated stationary signal fulfilling the inequation (7-1), while the remaining term $\boldsymbol{x}_n$ over-exceeds the a priori degree of accuracy defined by ϵ and should be related to the non–stationary corrupting process.
Thus, given an observation sample set of assumable second order time–series, $\boldsymbol{x} = \{x(t)\}$, with a discrete sampling period Δt. Its stochastic properties are considered stationary if a

given estimator $\Theta \subset \mathbb{R}$ on the observed measures does not change when shifted in time trough a length T, hence to measure the stochastic stationary properties of the given input process, we define the input stochastic feature matrix as an enhanced subspace based representation of the given process

Definition 7.1.3. (*Input stochastic feature matrix*): Let $x(t) \subset \mathcal{L}_2(T)$ denote a given Hilbert random signal such that the following narrow–band decomposition takes place: $\boldsymbol{x} \overset{\mathcal{T}}{\mapsto} \{\boldsymbol{y}_n : n = 1, \ldots, n_S\}$. Therefore, we call the training space a set of n_S stochastic feature vectors $\{\boldsymbol{y}_n^{(m)} \subset \mathbb{R}^{1\times T}\}$, where $\boldsymbol{y}_n^{(m)} = \left[y_n^{(m)}(1), \ldots, y_n^{(m)}(T)\right]$, being $y_n^{(m)}(t)$ the n-th feature measured from a given m-th observation, with $m \in n_O$ and $n \in n_S$, at a concrete time instant t. Accordingly, the extracted feature set is arranged into each m-th input stochastic matrix $\boldsymbol{Y}^{(m)} \subset \mathbb{R}^{n_S\times T}$:

$$\boldsymbol{Y}^{(m)} = \begin{bmatrix} \boldsymbol{y}_1^{(m)} \\ \vdots \\ \boldsymbol{y}_n^{(m)} \\ \vdots \\ \boldsymbol{y}_{n_S}^{(m)} \end{bmatrix} = \begin{bmatrix} y_1^{(m)}(0), & \ldots, & y_1^{(m)}(t), & \ldots & y_1^{(m)}(T) \\ \vdots & & \vdots & & \vdots \\ y_n^{(m)}(0), & \ldots, & y_n^{(m)}(t), & \ldots & y_n^{(m)}(T) \\ \vdots & & \vdots & & \vdots \\ y_{n_S}^{(m)}(0), & \ldots, & y_{n_S}^{(m)}(t), & \ldots & y_{n_S}^{(m)}(T) \end{bmatrix} \tag{7-3}$$

In order to enhance the time-series representation within the subspace framework, we employ an adapted version of commonly known nonparametric Singular Spectrum Analysis (SSA) method as a concrete decomposition method, $\mathcal{T}$. This approach is commonly used for analyzing hidden structures of time-series and aims at decomposing the original signal into a sum of the small number of meaningful components [31]. Furthermore, SSA analyzes hidden relationships among segments of input data making it an suitable tool for characterizing non-stationary time-series [16]. In Eq. (7-3), the supplied SSA-based extraction of the stochastic feature set is detailed in Algorithm 3.

7.2 Stochastic Constraints within Subspace-Based Analysis

The linear projection matrix to which the input stochastic matrix holds the most stationary features is to be find by incorporating stationarity conditions under the projection space to this the weak stationary constraints are rewritten as follows,

Definition 7.2.1. *Stationary Constraints for Input Stochastic Feature Matrix* Given an input stochastic feature matrix $\boldsymbol{Y} = \left\{\boldsymbol{Y}^{(m)} : m \in n_o\right\}$ and the linear projection matrix

Algorithm 3 SSA-based extraction of the stochastic feature set

Input: Input time-series object set $\{\boldsymbol{y}^{(m)} : m \in n_0\}$, window lag $L \in T$, number of elementary matrices n_S.
Output: Stochastic feature set, $\{\boldsymbol{x}_n\}$.

1. **Input time-series decomposition**

– *Embedding*: Given a length window L, the trajectory matrix $\boldsymbol{H} \subset \mathbb{R}^{L\times K}$, is computed by mapping the input time-series into a sequence of $K = T - L$ lagged vectors, $\boldsymbol{\zeta}_i^{(m)}$, that is, $\boldsymbol{H} = [\boldsymbol{\zeta}_1^{(m)}, \ldots, \boldsymbol{\zeta}_i^{(m)}, \ldots, \boldsymbol{\zeta}_K^{(m)}]$, where $\boldsymbol{\zeta}_i^{(m)} = [y^{(m)}(i), \ldots, y^{(m)}(i + L - 1)]$.

– *Eigenvalue decomposition*: SVD matrix is carried out that is a sum of rank-1 matrices corresponding to the n-th eigentriple of the following non-stationary decomposition [16]:

$$\boldsymbol{H} = \sum_{n\in \mathrm{rank}(\boldsymbol{H})} \boldsymbol{H}_n^* \tag{7-4}$$

where $\boldsymbol{H}_n^* = h_n \boldsymbol{h}_n \boldsymbol{h'}_n^{\top}$, $\{h_n \subset \mathbb{R}^+\}$ is the singular value set, $\boldsymbol{h}_n$ and $\boldsymbol{h'}_n$ are the right and left decomposition singular vectors, respectively.

2. **Stochastic feature extraction**

– *Grouping*: Once expansion in Eq. (7-4) is carried out, the grouping procedure tends to partition the n_S elementary matrices into the same number of subsets (*eigentriple grouping*), being $n_S = \mathrm{rank}(\boldsymbol{H})$.

– *Diagonal averaging:* We map each resultant matrix of Eq. (7-4) into n_S additive components of the initial time-series, $\boldsymbol{x}_n^{(m)}$. Afterwards, each stochastic feature $\boldsymbol{x}_n^{(m)}$ is arranged as the n-th row vector of the needed m-th input stochastic matrix $\boldsymbol{X}^{(m)}$.

Note: The following condition must be fulfilled: $\boldsymbol{y}^{(m)} = \sum_{\forall n} \boldsymbol{x}_n^{(m)}$.

$\boldsymbol{\Phi} \subset \mathbb{R}^{n_s \times n_s}$, weak stationary definition is written as,

$$\boldsymbol{E}\{\Theta\left\{\boldsymbol{\Phi}^{\mathsf{T}}\boldsymbol{Y}^{(m)};t\right\}:\forall m\} = \boldsymbol{E}\{\Theta\left\{\boldsymbol{\Phi}^{\mathsf{T}}\boldsymbol{Y}^{(m)};t\right\}:\forall t\} \tag{7-5}$$

It is an impose condition then, that the time average of the given estimator $\Theta\{\cdot;t\}$ must coincide with the ensemble average. Thereby a group-wise stationarity constraints take place:

$$\boldsymbol{\Phi}\bar{\boldsymbol{\mu}}_{\boldsymbol{Y}^{(m)}} = \boldsymbol{\Phi}\tilde{\boldsymbol{\mu}}_{\boldsymbol{Y}}, \tag{7-6a}$$

$$\boldsymbol{\Phi}\overline{\boldsymbol{\Sigma}}_{\boldsymbol{Y}^{(m)}}\boldsymbol{\Phi}^{\top} = \boldsymbol{\Phi}\widetilde{\boldsymbol{\Sigma}}_{\boldsymbol{Y}}\boldsymbol{\Phi}^{\top} \tag{7-6b}$$

where $\bar{\boldsymbol{\mu}}_{\boldsymbol{Y}^{(m)}} = \boldsymbol{E}\{\boldsymbol{y}_n^{(m)}:\forall t\}$, with $\bar{\boldsymbol{\mu}}_{\boldsymbol{Y}^{(m)}} \subset \mathbb{R}^{n_S\times 1}$, is the $n-$th stochastic feature vector mean along the time of the m-th observation, $\overline{\boldsymbol{\Sigma}}_{\boldsymbol{Y}^{(m)}}(n,l) = \left(\boldsymbol{y}_n^{(m)\top}\boldsymbol{y}_l^{(m)}\right)\delta(n-l)$, $\{n,l\} = 1\ldots n_O$, with $\overline{\boldsymbol{\Sigma}}_{\boldsymbol{X}^{(m)}} \subset \mathbb{R}^{n_S\times n_S}$, is the diagonal variance matrix of each stochastic feature of the $m-$th observation. In the same way, $\tilde{\boldsymbol{\mu}}_{\boldsymbol{Y}} = \boldsymbol{E}\{\bar{\boldsymbol{\mu}}_{\boldsymbol{Y}^{(m)}}:\forall m\}$, with $\tilde{\boldsymbol{\mu}}_{\boldsymbol{Y}} \subset \mathbb{R}^{n_S\times 1}$, is the ensemble mean value (i.e., the average computed over all the observations), and $\widetilde{\boldsymbol{\Sigma}}_{\boldsymbol{Y}} = \boldsymbol{E}\{\overline{\boldsymbol{\Sigma}}_{\boldsymbol{Y}^{(m)}}:\forall m\}$, with $\widetilde{\boldsymbol{\Sigma}}_{\boldsymbol{Y}} \subset \mathbb{R}^{n_S\times n_S}$, is the ensemble variance diagonal matrix.

Generally, estimation optimal projection matrix Φ that fulfills the constraints in equation (7-5) remains an open issue, since non–stationary projection is not always achievable. However, if a time-constant covariance structure of the stationary and non–stationary process is assumed, a generalized eigenvalue problem based solution can be posed as developed in [20]. The SSA optimization problem results in an linear projection matrix to which projected signals are as stationary as possible.

Particularly, an objective function based on the Kullback-Lieber divergence cadn be derive from the stationarity constraints in equation (7-3), aiming to minimize the distance of the mean and covariance of the stochastic features among observations. Additionally, in order to hold an orthonormal projection matrix, the matrix condition $\boldsymbol{\Phi}\tilde{\Sigma}\boldsymbol{\Phi}^{\mathsf{T}} = \boldsymbol{I}_{n_S}$, to this, the following generalized Rayleigh quotient is cast,

$$\begin{cases} \underset{\boldsymbol{\Phi}}{\operatorname{argmin}}\left\{\operatorname{tr}\{\boldsymbol{\Phi}\boldsymbol{\Xi}\boldsymbol{\Phi}^{\mathsf{T}}\}\right\} \\ \text{s. t.: } \{\boldsymbol{\Phi}\widetilde{\boldsymbol{\Sigma}}_{\boldsymbol{Y}}\boldsymbol{\Phi}^{\mathsf{T}}\} = \boldsymbol{I}_{n_S}, \end{cases} \tag{7-7}$$

And taking into account group–wise stationarity constraints given in Eqs. (7-6a) and (7-6b), we propose to compute the matrix $\boldsymbol{\Xi}$ as follows:

$$\boldsymbol{\Xi} = \boldsymbol{E}\{\bar{\boldsymbol{\mu}}_{\boldsymbol{Y}^{(m)}}\bar{\boldsymbol{\mu}}_{\boldsymbol{Y}^{(m)}}^{\mathsf{T}} + 2\overline{\boldsymbol{\Sigma}}_{\boldsymbol{Y}^{(m)}}\widetilde{\boldsymbol{\Sigma}}_{\boldsymbol{Y}}^{-1}\overline{\boldsymbol{\Sigma}}_{\boldsymbol{Y}^{(m)}}^{\mathsf{T}} : \forall m\} - \tilde{\boldsymbol{\mu}}_{\boldsymbol{Y}}\tilde{\boldsymbol{\mu}}_{\boldsymbol{Y}}^{\mathsf{T}} - 2\widetilde{\boldsymbol{\Sigma}}_{\boldsymbol{Y}} \tag{7-8}$$

Here, the matrix $\boldsymbol{\Xi} \subset \mathbb{R}^{n_S \times n_S}$ can be interpreted as the variance of the mean and covariance along all the observations. Thus, objective function optimization in Eqs. (7-7) and (7-8) aims to preserve the introduced constraints over the entire stochastic feature set through all the observations. As a result, the optimization problem stated in Eq. (7-7) can be solved in the form:

$$\boldsymbol{\Xi}\boldsymbol{\Phi} = \boldsymbol{\nu}\widetilde{\boldsymbol{\Sigma}}_{\boldsymbol{Y}}\boldsymbol{\Phi}, \tag{7-9}$$

where $\boldsymbol{\nu} \subset \mathbb{R}^{n_S \times n_S}$ is the set of generalized eigenvalues that along with the projecting matrix eigenvectors are further used to separate the stochastic feature set fulfilling the restriction of Eq. (7-5).

7.3 Principal Component Analysis and Projected Contribution to input Stationarity

As stated before, upon stationary constraints (7-6a) and (7-6b), optimization solution in Eq. (7-9) provides a set of projected stochastic features $\{\boldsymbol{\chi}_n^{(m)}(t) \subset \mathbb{R}^{1\times T} : n \in n_S\}$, that are as stationary as possible. Therefore, we must compute the contribution to input stationarity of the stochastic feature set, which is projected through the mapping matrix,

$$\boldsymbol{\Phi}^{\top}\boldsymbol{Y}^{(m)} = \begin{bmatrix} \boldsymbol{\chi}_1^{(m)}(t) \\ \vdots \\ \boldsymbol{\chi}_{n_S}^{(m)}(t) \end{bmatrix}$$

In particular, we measure contribution of each mapped feature, $\boldsymbol{\chi}_n(t)$, over all observations an at every time moment $t \in T$, in terms of projected variability grounded in the non-stationary reduction dimension method developed in [31]. To this purpose, we introduce an additional variability measure of projected stochastic feature, defined as follows:

$$\boldsymbol{E}\{var\{\{\boldsymbol{\Phi}^{\mathsf{T}}\boldsymbol{Y}^{(m)}; t\} : \forall m \in n_O\}\} = g_n(t) \tag{7-10}$$

To this, the following variability measure definition is formally given:

Definition 7.3.1. (*Variability measure of projected stochastic feature*): Contribution of the n-th projected stochastic feature, $g_n \subset \mathbb{R}^{+}$, is given in terms of following measure of

dispersion:

$$g_n = 1 - \sigma' \left\{ g_n\left(t\right) : \forall t \in T_n \right\}, \; T_n \in [(n-1)T + 1, nT]. \tag{7-11}$$

where $g_n\left(t\right) = \boldsymbol{E}\{\lambda_i^2 \varsigma_i^2\left(t\right) : \forall i\}$ is the amount of variability of the n-th projected stochastic feature, captured at the moment t, $\{\lambda_i \subset \mathbb{R}^+\}$ and $\{\boldsymbol{\varsigma}_i \subset \mathbb{R}^{1 \times n_S T}\}$ are the eigenvalue and eigenvector sets computed respectively from the SVD of $\boldsymbol{X}_{\boldsymbol{\Phi}}$. Notation $\boldsymbol{\varsigma}_i\left(t\right)$ stands for i-th eigenvector value at $t \in n_S T$. Besides, $\sigma' \subset \mathbb{R}[0, 1]$ is a normalized standard deviation. We denote the projected feature set by $\boldsymbol{Y}_{\boldsymbol{\Phi}} = \left[\vec{\boldsymbol{\chi}}^{(1)\mathsf{T}} \cdots \vec{\boldsymbol{\chi}}^{(m)\mathsf{T}} \cdots \vec{\boldsymbol{\chi}}^{(n_O)\mathsf{T}}\right]^{\mathsf{T}}$, with $\boldsymbol{Y}_{\boldsymbol{\Phi}} \subset \mathbb{R}^{n_O \times n_S T}$, where each row corresponds to the vectored form of the $m-$th observation mapping matrix $\vec{\boldsymbol{\chi}}^{(m)} = \mathrm{vec}(\boldsymbol{\Phi}^{\mathsf{T}} \boldsymbol{X}^{(m)})^{\mathsf{T}}$.

Derivation of the variability distance of Eq. (7-11) is detailed in A, which basically measures the inertia of each column-wise feature [1]. Once the variability measure of projected space is computed, we reconstruct the stationary input matrix by weighting the contribution of each projected feature, as follows:

$$\boldsymbol{Y}_s^{(m)} = \boldsymbol{\Phi}\, \mathrm{Diag}(\boldsymbol{g}) \boldsymbol{\Phi}^{\mathsf{T}} \boldsymbol{Y}^{(m)}, \tag{7-12}$$

where weighting vector $\boldsymbol{g} = [g_1 \ldots, g_{n_S}] \subset \mathbb{R}^{n_S \times 1}$ comprises contribution of the projected stochastic feature set, given in terms of its variability. Afterwards, stationary components of each stochastic feature are extracted. Thus, by considering the bijective property of the SSA-based decomposition, stationary composite signal of input time-series can be reconstructed by the inverse transform, that is, $\boldsymbol{Y}_s^{(m)} \overset{\mathcal{T}^{-1}}{\mapsto} \hat{\boldsymbol{x}}_s^{(m)}$. Lastly, estimated non-stationary residual signal of input time-series, according to Eq. (7-2), can be obtained as:

$$\hat{\boldsymbol{x}}_n^{(m)} = \boldsymbol{x}^{(m)} - \hat{\boldsymbol{x}}_s^{(m)} \tag{7-13}$$

8. Validation Results

Quality of separating filtering task if performed by using an distance-based measure $\mathrm{d}\,(\boldsymbol{x}_s)\,\hat{\boldsymbol{x}}_s \subset \mathbb{R}^+$, quantifying similarity that expresses the degree of statistical dependency/independency between the desired stationary signal $\boldsymbol{x}_s$ and its estimate $\hat{\boldsymbol{y}}_s$ throughout the considered time interval $t \in T$. To validate the proposed filtering extraction methodology, simulated synthetic time series and real data are used, whereas the first case the real stationary component $\boldsymbol{x}_s$ is known, correlation coefficient is used as proximity estimation measured:

$$\rho(\boldsymbol{x}_s, \hat{\boldsymbol{x}}_s) = cov\{\boldsymbol{x}_s, \hat{\boldsymbol{x}}_s\}/\sigma_{\boldsymbol{x}_s}\sigma_{\hat{\boldsymbol{x}}_s},\ \rho(\boldsymbol{x}_s, \hat{\boldsymbol{x}}_s) \subset \mathbb{R}[-1, 1] \tag{8-1}$$

where $\sigma_{\boldsymbol{x}_s}$ and $\sigma_{\hat{\boldsymbol{x}}_s}$ are the standard deviation values of $\boldsymbol{x}_s$ and $\hat{\boldsymbol{x}}_s$, respectively. The case $|\rho| \to 1$ occurs when $\boldsymbol{x}_s$ and its estimate $\hat{\boldsymbol{x}}_s$ tend to perfectly coincide, otherwise, $|\rho| \to 0$.
In case of real data, the true stationary composite signal is not available and to quantify the degree of signal separation indirect inference methods should be used. Particularly, since temporal fluctuation in spectral components of the signal time-frequency representation $\boldsymbol{S}_x\,(\omega, t)$ points out on signal non-stationarity, the degree of stationarity averaged over all frequency domain is taken as the quality measure [39]:

$$d_{\hat{\boldsymbol{x}}} = \boldsymbol{E}\{\boldsymbol{d}\,(\hat{\boldsymbol{x}}; \omega) : \forall \omega \in \Omega\},\ d_{\hat{\boldsymbol{x}}} \subset \mathbb{R}^+ \tag{8-2}$$

where the vector distance $\boldsymbol{d}(\hat{\boldsymbol{x}}; \omega)$ is computed through the marginal frequency distribution along the frequency domain, that is, $\boldsymbol{d}\,(\hat{\boldsymbol{x}}; \omega) = \boldsymbol{E}^2\{1 - \boldsymbol{S}_{\hat{\boldsymbol{x}}}\,(\omega, t)/\overline{\boldsymbol{S}}_{\hat{\boldsymbol{x}}}\,(\omega) : \forall \tau \in T\}$ being $\overline{\boldsymbol{S}}_{\hat{\boldsymbol{x}}}\,(\omega) = \boldsymbol{E}\{\boldsymbol{S}_{\hat{\boldsymbol{x}}}\,(\omega, t)\,, \forall t \in T\}$. In case of stationary signal, the degree of stationarity yields low and constant values along frequency spectrum, while it increases both its value and variability in the presence of non-stationarity.
The proposed methodology, termed as *Subspace-based Separation with Stationarity Constraints (SS-SC)* from now on , between non-stationary and stationary signals is carried out according to the scheme presented in Figure **8-1**. The methodology accomplishes the following three subsequential stages: *i*) Extraction of stochastic feature set, *ii*) Subspace-based projection under proposed stochastic constraints, and *iii*) Evaluation of projected contribution to input stationarity. SS-SC methodology is evaluated on simulated data as well as on

Motor/Imagery-EEG real data for the purpose of decomposing EEG activity.

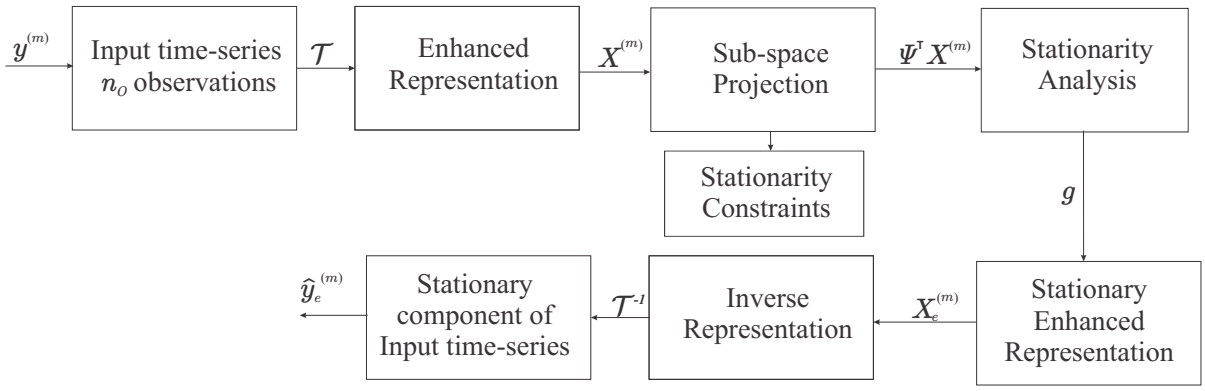


Figure 8-1: Testing experimental setup scheme of the proposed SS-SC methodology of separation between non-stationary and stationary signals based on subspace analysis

For the sake of comparison, the SS-SC is contrasted against other two similar projection approaches. Particularly, the *Second Order Non-Stationary Sources* approach (SONS) [9] and the *Analytic Stationary Subspace Analysis* (ASSA) [20] are employed. Besides, along with the proposed variability measure of projected stochastic feature, in the sequel noted as $\boldsymbol{g_\chi} \subset \mathbb{R}[0,1]$, another baseline measure is also considered: $\boldsymbol{g_x} \subset \mathbb{R}$ that is computed as the standard deviation of projected components. In case of ASSA projection, the measure $\boldsymbol{g_x} \in \{0,1\}$ is obtained by setting the weight of the eigenvectors associated with the smallest eigenvalues as 1 and 0 with the remaining values. For SONS projection, since there is no an eigenvalue set, the weights are set as the normalized standard deviation of projected components. Both considered variability measures hold dimension $n_S \times 1$.

As a result, six methods are compared for separation between non-stationary and stationary signals appraising considered projection approaches as well as variability measures, as summarized in Table 8-1. Separation quality is evaluated in two different scenarios: *i*) $n_O = 1$, i.e., a single observation of a one-dimensional time–series, and *ii*) $n_O > 1$, when several observations of one-dimensional time-series are available. To ensure the same conditions of evaluation, in the former scenario, methods one to four have to be carried along with the enhancement representation stage, due to the inability of the ASSA and SONS subspace projections to deal with one-dimensional time-series.

Notation	*Subspace projection*	*Variability measure*
Method 1	ASSA	$\boldsymbol{g_x}$
Method 2	ASSA	$\boldsymbol{g_\chi}$
Method 3	SONS	$\boldsymbol{g_x}$
Method 4	SONS	$\boldsymbol{g_\chi}$
Method 5	SS-SC	$\boldsymbol{g_x}$
Method 6	SS-SC	$\boldsymbol{g_\chi}$

Table 8-1: Summary of compared subspace-based methods in separation between non-stationary and stationary signals

8.1 Numerical results on simulated data

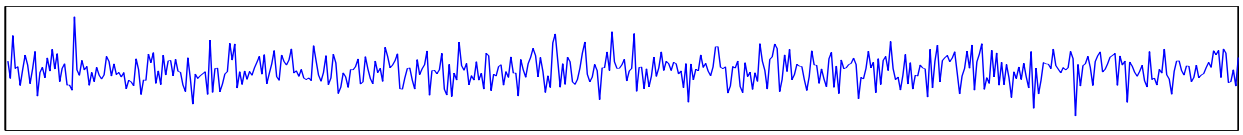
Dataset description According to the model of Eq. (7-2), all observed time–series are generated as linear superposition of simulated stationary and non-stationary sources. As proposed in [20], the stationary signals are generated by one of the following two models: either randomly drawn from a normal distribution, i.e., i.i.d. Gaussian $\mathcal{N}(\boldsymbol{\mu}_s, \boldsymbol{\Sigma}_s)$, where $\boldsymbol{\mu}_s, \boldsymbol{\Sigma}_s$ are respectively the mean and covariance matrices of the normal distribution, or by an autoregressive moving average model (particularly, a $(3,3)$ model is used). Parameters of either representation are randomly sampled from the Normal distribution $\mathcal{N}(\mathbf{0}, \mathbf{1})$. The non-stationary signals are formed by one of the following models: an i.i.d. Gaussian model, $\mathcal{N}(\boldsymbol{\mu}_n, \boldsymbol{\Sigma}_n)$, that includes some induced abrupt changes, whose number is also random ranging from 10 to 40 within the considered 500 sampling length or from chaotic Lorenz attractor with parameters also randomly sampled from the Normal distribution. Figure **8-2** shows an example of the generated signals: stationary (see Fig. 8.2(a) and Fig. 8.2(b)) and non–stationary (Fig. 8.2(c) and Fig. 8.2(d).

For provided SSA-based enhancement of generated time-series set, the number of samples is set to be $T = 500$ while the lag parameter is empirically fixed as $L = 50$.

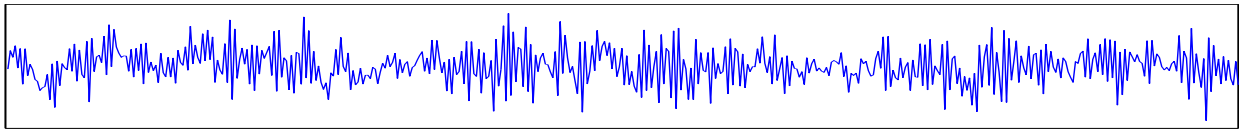
8.1.1 Performed separation of stationary signals

Quality of signal separation is performed according to the correlation index, $\rho(\boldsymbol{y}_e, \hat{\boldsymbol{y}}_e)$, given in Eq. (8-1) and introduced in the case of simulated data. Also, influence of the non-stationary signal power on estimated quality is considered, varying the following power ratio: $\kappa = \|\hat{\boldsymbol{y}}_e\| / \|\hat{\boldsymbol{y}}_n\|$.

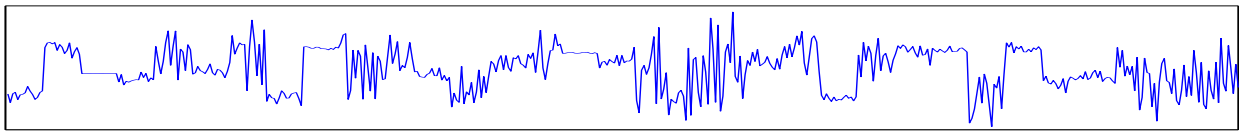
At the beginning, stationary separation is carried out over a single observation clipped from one-dimensional time–series. Figure 8.3(a) shows Correlation index, $\rho(\boldsymbol{y}_e, \hat{\boldsymbol{y}}_e)$, that is



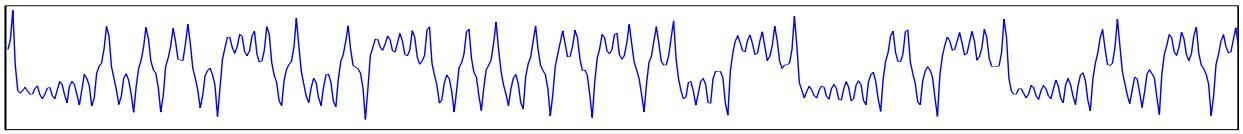
(a) Stationary, Gaussian component.



(b) Stationary, ARMA component.



(c) Non-stationary, Gaussian-distributed abrupt changes.



(d) Non-stationary, Chaotic Lorenz attractor.

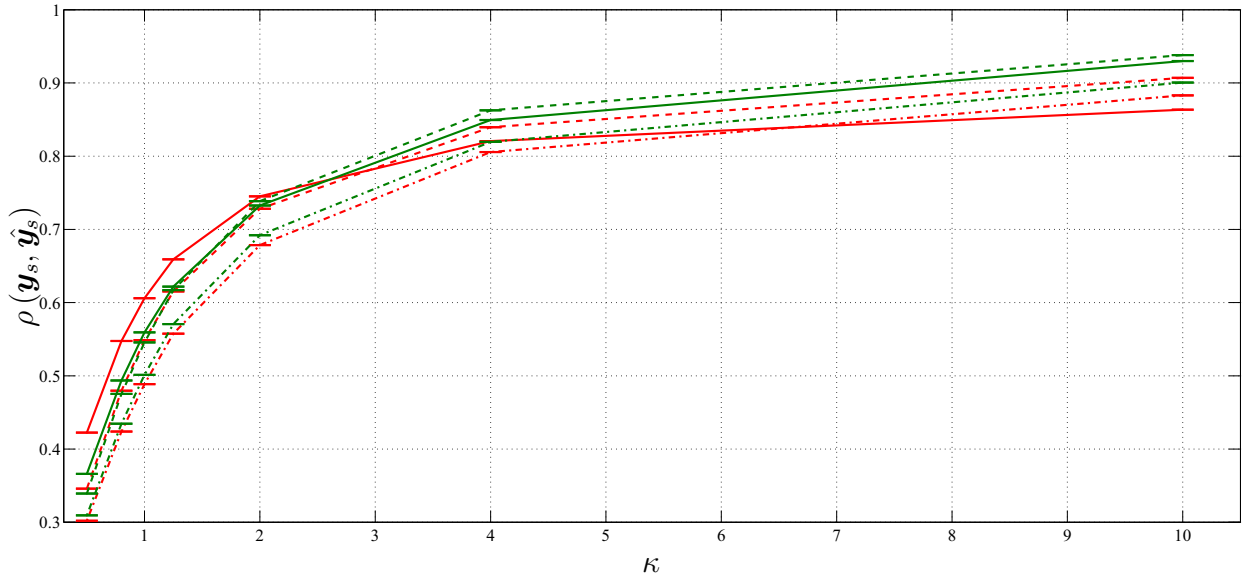
Figure 8-2: Examples of generated stationary and non-stationary real-valued time series.

computed using the Monte Carlo experiment with 600 repetitions, when the power ratio κ ranges within interval $[0-10]$. As seen, The ASSA projection overperforms slightly, though not dramatically, both the SS-SC and SONS approaches regardless of the used variability measure. However, the introduction of the measure $\boldsymbol{g}_{\boldsymbol{\chi}}$ improves all projections.

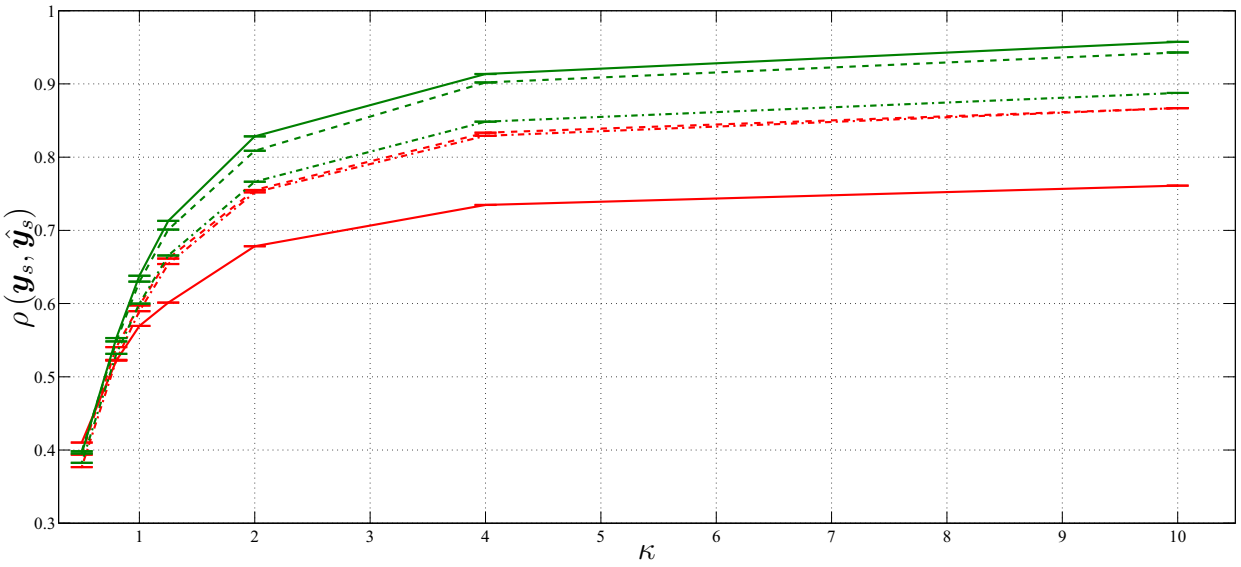
In case of multiple observations, the performed correlation index is computed also in the Monte Carlo setup over 48 repetitions for different values of κ, where each set contains 30 non-consecutive observations. As seen in Figure 8.3(b), although SS-SC fails if combining with the $\boldsymbol{g}_x$ measure, this projection together with the projected variability measure $\boldsymbol{g}_{\chi}$ clearly gets the best quality. So, one may infer that SS-SC takes better advantage of more information because of the higher number of observations. Therefore, in the reminder of this experimental part, we will use only the projected variability measure $\boldsymbol{g}_{\chi}$.

8.1.2 Estimated computational burden

In case of the single observation, SONS algorithm requires for a whitening procedure as well as time-windowed computations, increasing remarkably the processing time. In contrast, ASSA only demands of time-windowed computations, decreasing the computational burden. However, the SS-SC uses a simple weighted stationary projection, making it the fastest. It



(a) one single observation, $n_O = 1$



(b) multiple observations, $n_O = 30$

Figure 8-3: Computed correlation index, $\rho(\boldsymbol{y}_e, \hat{\boldsymbol{y}}_e)$, between original and corresponding estimated stationary signals of simulated data. Red lines display $\boldsymbol{g}_x$ while green lines the $\boldsymbol{g}_\chi$ measure. Dashed-dot lines stand for SONS, dashed lines for ASSA, and solid for SS-SC.

must be quoted that in case of multiple observations, methods 5 and 6 include an additional stage performing the stochastic feature extraction, increasing computational burden unlike remaining methods in which sub-space projection is computed directly. To have an idea, Table 8-2 compares the consumed time over one observation and over a set of 30 observations, using studied projection approaches. The variability influence calculation is not taken into consideration because of its relative very low time demand.

n_O	*SONS*	*ASSA*	*SS-SC*
1	1.77 ± 0.073	0.078 ± 0.005	$\mathbf{0.057 \pm 0.003}$
30	0.68 ± 0.84	$\mathbf{0.02 \pm 0.063}$	1.6 ± 0.033

Table 8-2: Computation time $[s]$ of stationary signal separation on dependence of studied projection approach.

8.2 Application on EEG Motor Movement/Imagery data

8.2.1 Brain computer interface data

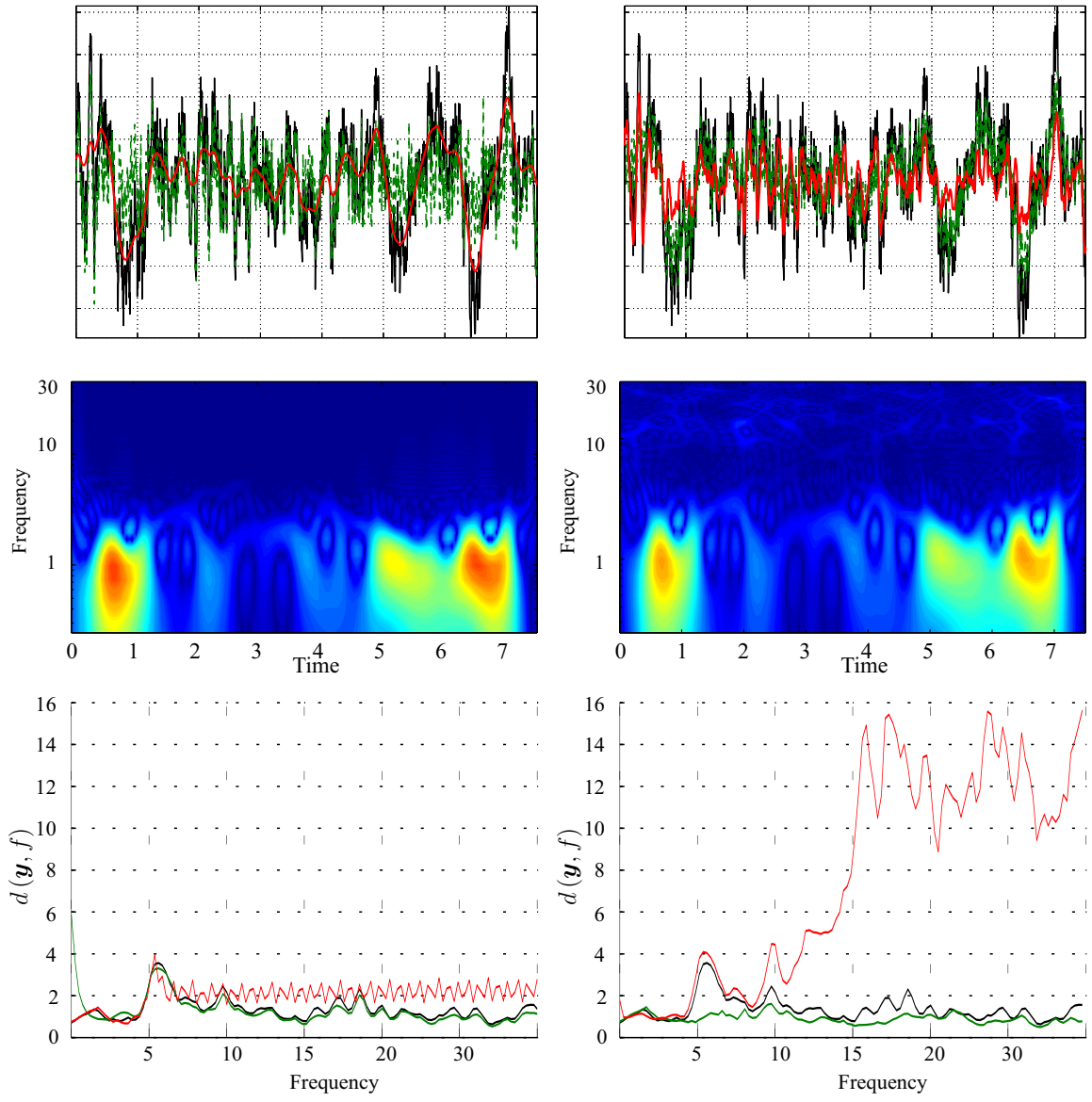
As real data, we employ brain computer interface cue-based paradigm dataset that is used for motor-imagery classification and provided by the Graz University of Technology [37]. EEG data hold recordings of 9 subjects who were asked to perform one of the following four tasks, at a given cue triggered time: the movement of the left hand (class 1), right hand (class 2), both feet (class 3), and tongue (class 4). Each subject recording session includes approximately 5 minutes of reference data (i.e., free of triggers or movements) to estimate electrooculogram influence. To this, the reference data are divided into 3 blocks: *i*) two minutes with eyes open, *ii*) two minutes with eyes closed, and *iii*) one minute with eye movements. Once the subjects were asked to sit in a comfortable armchair in front of a computer screen, movement trials started after reference recording and were separated by short breaks. At the beginning of each trial, a fixation cross appeared in the black screen and short duration acoustic warning tone was also presented. After two seconds, an arrow cue appeared indicating one of the four class directions (left,right, down, or up). The cross stayed on the screen for approximately $3\,s$ which triggered then end of the desired movement. Database supplies labeling of time instants, duration of each reference block, and cue-based movement. Acquired signals were recorded at sampling rate of $250\,Hz$ using 22 Ag/AgCl electrodes. Recordings were bandpass-filtered within a $[0.5 - 100]\,Hz$ bandwidth.
We consider only the subject #1 in the three-class mode, namely, reference state (labeled as *class 1*), left hand movement (*class 2*), and right hand (*class 3*). Besides, all observations are

taken from several trials of the selected Cz EEG channel that is related to the sensorimotor cortex. Thus, n_O trials of one dimensional time-series are obtained. As the preprocessing stage, enhanced representation of used Cz channel is accomplished with the lag parameter set to $L = 60$ and bounded by the empirical rule of thumb $L \geq f_s/f_m$, where f_s and f_m are the sampling frequency and minimum desired frequency of representation, respectively. So, for the motor-imagery task, L is expected to be around the minimum frequency of the θ band (i.e., $4\,Hz$).

8.2.2 Performed signal separation over EEG data

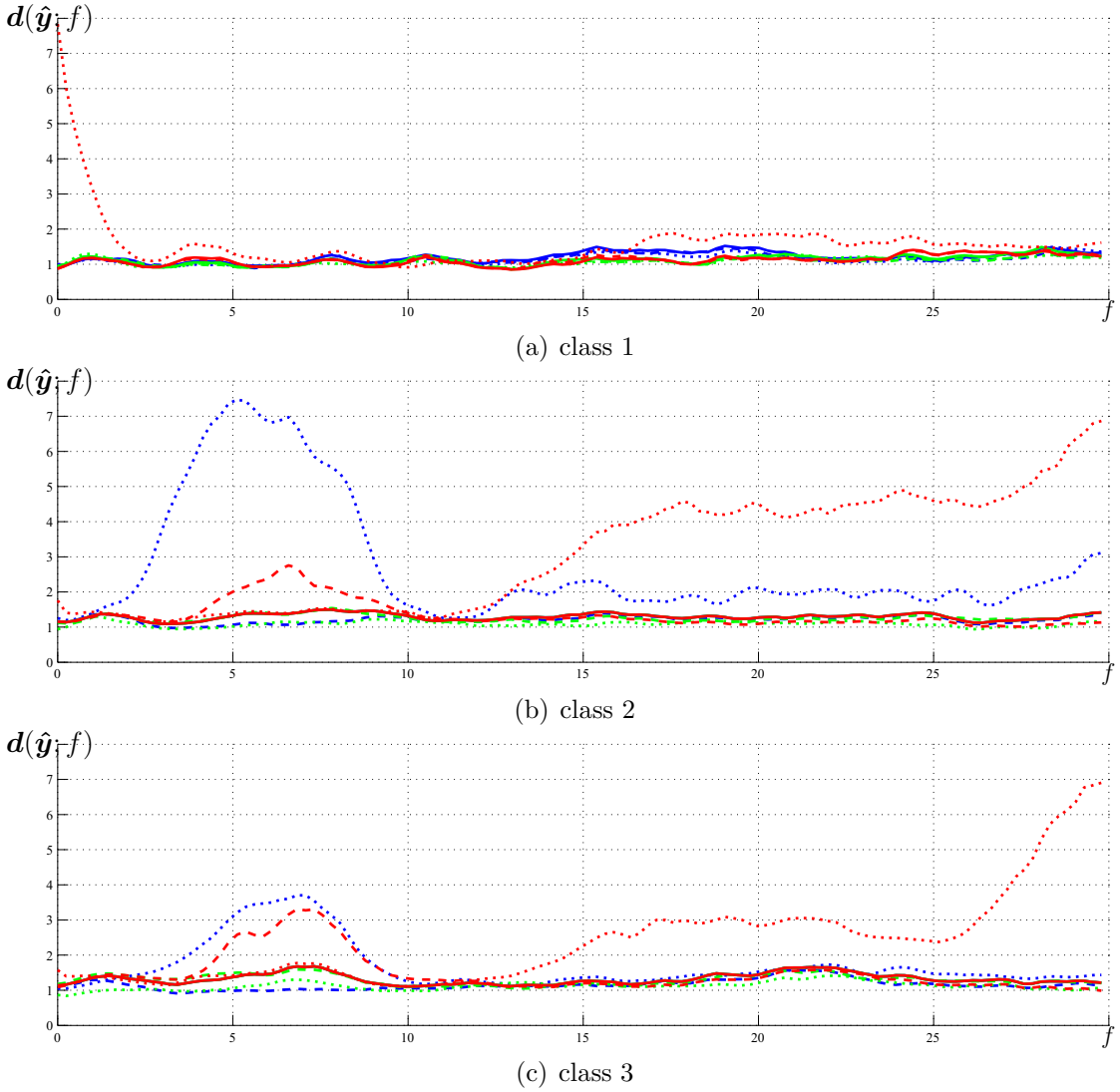
A visual illustration of accomplished signal separation using SS-SC is given in Figure **8-4**. The top row shows the oscillograms of an input time–series at hand, as well as separated stationary and non-stationary signals. Separation is performed in both scenarios: over each trial individually (single observation mode, namely, trial 19 belonging to the class 1) and over 49 trials per class (multiple observations). As explained above, the quality of separating filtration over real data is performed using the degree of stationarity, (see Eq. (8-2)), where as time-frequency representation the smoothed pseudo-Wigner-Ville distribution (PSWVD) is computed. As seen, the computed time-frequency representations of separated stationary (second row) and non-stationary (third row) signals evidence that the separation filtering task may still work under the assumption that the time-frequency distributions of considered input time–series overlap. The bottom row displays the estimated degree of stationarity through time, making clear how the stationary separation method improves its effectiveness as the observation number increases in both considered scenarios. However, PSWVD calculation may supply collateral cross-terms distorting frequency boundaries, which in turn results in bursting peaks of the boundary $d_{\boldsymbol{y}}$ values.

Figure **8-5** shows computed quality of separation for all three projections under comparison (Green lines represent SONS, blue – ASSA, and red–SS-SC). Our basic assumption is that any movement of either hand must increase non-stationarity of the measured Cz EEG channel. Hence, in case of reference class 1, the degree of stationarity should be mostly similar for both input EEG and separated stationary signals. On the contrary, in case of classes 2 or 3, quality values through frequency should be unalike as possible for both input and separated non-stationary signals. As seen in Figure 8.5(a) showing the estimated values of $\boldsymbol{d}(\hat{\boldsymbol{y}}_s; f)$, of class 1, input (solid lines) and separated stationary (dashed lines) signals keep very similar trajectory, while the contribution of non-stationary signal (dashed-dot lines) is very discrete. Overall, the considered projections perform similar quality. For both motor tasks, Figures 8.5(b) and 8.5(c) shows that separated non-stationary signal has pronounced



(a) $n_O = 1$ (b) $n_O = 49$

Figure 8-4: Exemplary of computed time-frequency representations of stationary (middle raw) and non-stationary (bottom raw) signals using SS-SC from *Cz* EEG channel (class 2). Degree of Stationarity: (■ Original, ■ Stationary, and ■ Non-Stationary signal)



(a) class 1

(b) class 2

(c) class 3

Figure 8-5: Computed values of Degree of stationarity, $\boldsymbol{d}(\hat{\boldsymbol{y}}; f)$, for separated stationary and non-stationary signals from *Cz* Channel. Green lines represent SONS, blue – ASSA, and red – SS-SC. Solid lines represent input EEG activity, dashed lines – separated stationary signals, and dashed-dot lines – non-stationary signals.

contribution, for ASSA and SS-SC (but not for SONS). Besides, non-stationarity of electrical activity is focalized within two frequency sub–bands: $(2-10)\,Hz$ and $>13\,Hz$, at least, for SS-SC. ASSA projection detects just the first sub–band, while SONS does not at all. This finding reflects the neural activity of the motor/imagery task in the following spectral bands: θ $(4-7)$, α $(7-14)$, and β $(15-30)Hz$, as discussed in [29].

9. Discussion and conclusions

The main goal of the present chapter is to supply a complex of separation filtering algorithms between stationary and non-stationary signals from a given time-series. For this purpose, we suggest the Subspace-based Separation with Stationary Constraints, which lies in the hypothesis that enhancing the input signal representation together with additional stationary constraint over the projected space should improve quality of the separation filtering task. The obtained results evidence the following aspects to take into consideration:

i) Introduced time-series enhancement for extracting the set of time-variant features should be regarded as an important factor for highlighting stochastic signal structure. To this end, features are extracted using singular spectrum analysis that allows analyzing relationship among segments of the input time-series. Nevertheless, other dynamic decomposition techniques may be used providing their bijective property, i.e., there exists an inverse transform that allows returning to the input space. Among others, some time-frequency/scale representations or Empirical Mode Decomposition are to be further contemplated.

ii) Stationary signal separation is carried out by imposing stochastic constraints to the matrix projection. The main goal behind subspace projection is to reveal the simpler structure that underlies it, besides of mitigating noise influence on the input signal. As the baseline projection, the SONS approach is considered that diagonalises simultaneously sample covariance matrices obtained from segments of the input time-series. Yet, SONS approach requires solving much computationally expensive non–convex optimization procedures. To overcome this issue, the ASSA projection is also considered that embraces an objective function based on the Kullback-Leibler divergence to compute both the mean and covariance through an introduced set of time segments. However, to improve quality of signal separation, we propose the SS-SC that uses objective function aiming to minimize the distance of the mean and covariance but across the entire observation ensemble.

iii) Along with the assumption usually made that the non-stationarities are measurable in the first two statistical moments, we also introduce evaluation of how the imposed

stationary constraints hold over the projected space. In practice, the projected space is empirically truncated into a lower dimension. But performed separation may degrade on depend on the used heuristical reduction (not mentioning the difficult to fully automate such a procedure). To cope with this situation, we make clear how each projected feature contributes to the stationarity of the input time-series. As a result, separating filtering task grounds on those reconstructed features better fulfilling the imposed stochastic constraints.

iv) Six methods are compared for separation between non-stationary and stationary signals appraising three projection approaches for two variability measures. Because of practical considerations, besides of convectional multi-observation case, separation quality is evaluated in case of one-dimensional time-series single observation. Obtained numerical results on simulated data show that every method may reach similar separation quality, though, the proposed variability measure improves performed separation, as seen in Figure **8-3**. In turn, carried out validation on real brain computer interface data shows that the proposed SS-SC takes better advantage of more information when having higher number of observations and it is the only method able to separate electrical activity focalized within two frequency sub–bands, while ASSA method discover just the first one, and SONS does not at all.

v) In terms of computational burden, SS-SC is the fastest when performing the single observation task, since used projection allows computing projected space without accomplishing the partition procedure of input time-series. Still, the needed additional stage performing stochastic feature extraction that must be carried out for each observation makes the SS-SC the most time consuming method. Therefore, proposed separating quality improvement is paid with a reasonable amount of increased computational cost, as seen in Table **8-2**.

As the concluding remarks, this work proposes the *Subspace-based Separation with Stationary Constraints* to discriminate between stationary and non-stationary signals. The method lies in the hypothesis that enhancing the input signal representation and introducing an additional stationary constraint over the projected space should improved quality of the separation filtering task. Besides, to evaluate how stationary restriction holds over the optimization solution within a given projection framework, we introduce a variability measure of projected stochastic feature. Carried out validation on both simulated and real data shows that SS-SC takes better advantage of more information from the higher number of observations. Nonetheless, proposed in this work separating quality improvement is paid with a

reasonable amount of increased computational cost.

Future work includes investigation of a more powerful projection that allows discriminating among strong focalized either in time or frequency non-stationary dynamics as well as a generalization of the stationary constraints for multiple channels and multivariate time-series. Application of the proposed separation method in more specific Motor/Imagery-EEG contexts is also to be further considered. Nevertheless, a more suitable quality measures should be developed to decompose electrical activity into meaningful components.

Part IV

Novel Recursive Stationary Signal Analysis from Time–Variant Feature Representation

In practice, stationary signal separation serves as a primary filtering task that permits distinguishing among different components based only upon the signal stochastic properties. Thereby, design and implementation of fast and recursive signal separation methodologies is usually required. In this section, a fast recursive method for stationary signal separation is presented, using recursive subspace analysis the computational burden of separation task is reduced by means of rank-one updates of the eigenvector structures of the subspaces. The stochastic decomposition of initial input data and its respective initial subspace representation serves as base model for stationary signal separation, however, as non–stationary processes appear, the recursive nature of the representation, updates at each sample step the stationary subspace. Additionally, previously introduced variability measure is also updated within the new window signal representation, thus process structure variation is taken into consideration sequentially at each stage.

Recalling the signal separation problem of Definition 7.1.2, the recursive nature of the proposed scheme, poses as an online solution for stationary signal separation analysis, thus equation (7-2) is rewritten as,

$$\mathcal{M}\left\{\boldsymbol{x}\left(t\right),\epsilon,\boldsymbol{\Psi}^{t^{-}}\right\}=\tilde{\boldsymbol{x}_{s}}(t)+\tilde{\boldsymbol{x}_{n}}\left(t\right) \tag{9-1}$$

where $\boldsymbol{\Psi}^{t^{-}}$ is the stationary subspace at time instant t^{-}.

10. Recursive Signal Stationary Analysis using Subspace Based Methods

Online signal separation is carried out by means of recursive decomposition of the input signal. However, this can be very time-consuming task if the whole whole analyzed signal is required in addition to the respective nature of the representation. In [27], a recursive SSA methodology is proposed using rank-one updates of the eigen–structure of the trajectory matrix (hankel matrix) $\boldsymbol{H} = \{\varsigma_k : k \in T\}$, with $\boldsymbol{H}_y \in \mathbb{R}^L \times K$, recall Algorithm 3.

In the static conventional SSA, the stochastic components of the input signal are elicited by projection of the trajectory matrix onto the eigen–space defined by the singular value decomposition of the covariance matrix $\boldsymbol{D} = cov\{\boldsymbol{H}_y, \boldsymbol{H}_y,\}$ of $\boldsymbol{H}_y$ [15]. Thus, posing an online signal analysis under conventional SSA would require updating the trajectory matrix whenever new data is collected, continuously increasing its size and consequently the computational burden of the eigen-decomposition. In turn, Recursive SSA (R-SSA) makes use of rank-one update for adjusting a given initial covariance matrix, holding a fixed size.

Particularly, for a given new sample, the sample covariance matrix is approximated by:

$$\boldsymbol{D}_k = \frac{k-1}{k}\boldsymbol{D}_{k-1} + \frac{1}{k}\varsigma_k\varsigma_k^{\mathsf{T}} \tag{10-1}$$

Thus, by letting $\boldsymbol{D}^{k-1} = \boldsymbol{U}^{k-1}\boldsymbol{\Sigma}\left(\boldsymbol{U}^{k-1}\right)^{\mathsf{T}}$, where $\boldsymbol{U} \in \mathbb{R}^{L\times L}$ and $\boldsymbol{\Sigma} \in \mathbb{R}^{L\times L}$ are the eigenvectors and the eignevalues matrices given by the SVD of matrix $\boldsymbol{D}$, at instant $k \in T$. The updated eigen–space can be computed under the perturbation analysis [27] for rank-one

update matrices as:

$$\begin{cases} p_{i,j}^k = \left(\boldsymbol{U}_i^k \tilde{\boldsymbol{y}}\left(k\right) - D_{i,j}^k\right) k^{-1}, & i = j = 1, \ldots, w \\ p_{i,j} = 0, & i \neq j \end{cases} \tag{10-2a}$$

$$\begin{cases} q_{i,j}^k = \frac{\boldsymbol{U}_i^k \tilde{\boldsymbol{y}}(k) \boldsymbol{U}_j^k \tilde{\boldsymbol{y}}(k)}{k\left(D_j^{k-1} - D_i^{k-1} + p_{j,j}^k - p_{i,i}^k\right)}, & i, j = 1, \ldots, w, i \neq j \\ q_{i,j} = 0, & i = j, \end{cases} \tag{10-2b}$$

being $\boldsymbol{I}$ a $w \times w$ identity matrix. As a result, R-SSA computes the sample stochastic components at new samples points k using the updated eigen-space at each sample. Algorithm 4 shows the general outline for computing R-SSA:

Algorithm 4 General Outline for R-SSA

1: Input time–series batch data $\boldsymbol{y} \in \mathbb{R}^{T_0}$, window lag size w for trajectory matrix.
2: Initialize $\boldsymbol{D}_0$ and $\boldsymbol{\Sigma}_0$ eigen-space.
3: **while** $y\left(k\right) : k > T_0$ **do**
4: Compute perturbation matrices $\boldsymbol{P}_k, \boldsymbol{Q}_k$ and update eigen-space $\boldsymbol{D}_k = \boldsymbol{P} + \boldsymbol{D}_{k-1}$ and $\boldsymbol{\Sigma}_k = \boldsymbol{\Sigma}_{k-1} + \boldsymbol{I}$ and normalize eigen-vector space norms.
5: Project sample trajectory matrix onto the eigen-space to elicit principal components $\boldsymbol{R}^k$ and calculate sample reconstructed stochastic components $\boldsymbol{X}^k$.

$$\boldsymbol{R}_i^k = \sum_{j=1}^{w} \tilde{\boldsymbol{y}}\left(k - j + 1\right) \boldsymbol{\Sigma}_j^k \tag{10-3a}$$

$$\boldsymbol{X}_i^k = \frac{1}{m_t} \sum_{i \in M} \sum_{j=l_t}^{u_t} \boldsymbol{r}_i^{k-j+1} \boldsymbol{\Sigma}_j^k \tag{10-3b}$$

with $m_k = \min\left(w, T_p\right)$, $l_k = \max\left(w, T_p\right)$ and $M = w + T_p - 1$
6: Output $\boldsymbol{X}^k$
7: **end while**

10.1 Non-stationary component extraction

As in section §8, signal stationary component separation is done in a sequential procedure, where for a given input signal, stochastic feature representation is used for generating a subspace whereas stationarity of projected components is ensured, additionally, the variability measure bias the signal reconstruction to those projected components wit highest stationarity. The newly introduction of recursive signal decomposition, thus poses a similar process, to which not only fewer parameterizations are required but also allows to maintain

computational simplicity. Figure 10-1 shows a general outline for an online signal stationary separation analysis of a given input data $y(t)$.

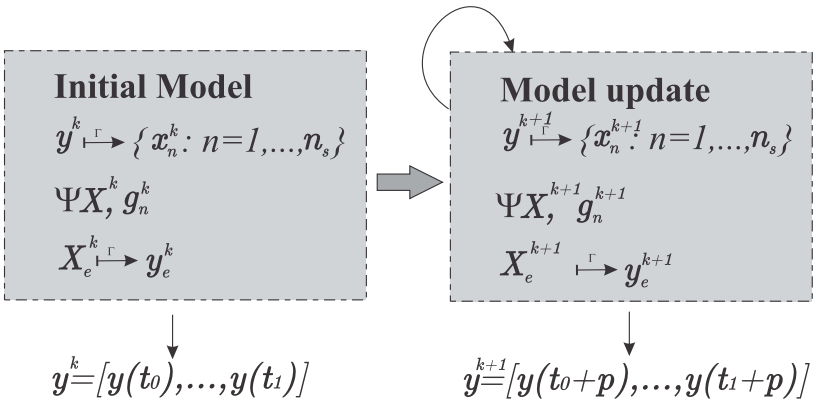


Figure 10-1: Schemating representation of proposed recursive approach of separating stationary components

10.1.1 Subspace representation with stationarity constraints under recursive analysis

As mentioned before, the main goal of conventional subspace analysis is to find a projected space $\boldsymbol{\Psi}^k \boldsymbol{X}^k$, bearing still as much information of the given process representation $\boldsymbol{X}^k$ as possible. As for time-variant recursive features are computed, the subspace structure must then change an adapt itself to the new introduced representation, thus equations (7-7) to (7-9) are to be rewritten in terms of the time-variant estimation of the input data,

$$\begin{cases} \underset{\boldsymbol{\Psi}^k}{\operatorname{argmin}} \left\{ \operatorname{tr}\{\boldsymbol{\Psi}^k \boldsymbol{\Xi}^k \boldsymbol{\Psi}^{k\top}\} \right\} \\ \text{s. t.: } \{\boldsymbol{\Psi}^k \widetilde{\boldsymbol{\Sigma}}_{\boldsymbol{X}}^k \boldsymbol{\Psi}^{k\top}\} = \boldsymbol{I}_{n_S}, \end{cases} \tag{10-4}$$

with $\boldsymbol{\Xi}^k$ given by:

$$\boldsymbol{\Xi}^k = \boldsymbol{E}\{\bar{\boldsymbol{\mu}}_{\boldsymbol{X}^k} \bar{\boldsymbol{\mu}}_{\boldsymbol{X}^k}^\top + 2\overline{\boldsymbol{\Sigma}}_{\boldsymbol{X}^k} \widetilde{\boldsymbol{\Sigma}}_{\boldsymbol{X}^k}^{-1} \overline{\boldsymbol{\Sigma}}_{\boldsymbol{X}^k}^\top\} - \tilde{\boldsymbol{\mu}}_{\boldsymbol{X}^k} \tilde{\boldsymbol{\mu}}_{\boldsymbol{X}^k}^\top - 2\widetilde{\boldsymbol{\Sigma}}_{\boldsymbol{X}^k} \tag{10-5}$$

where $\bar{\boldsymbol{\mu}}_{\boldsymbol{X}^k} = \boldsymbol{E}\{\boldsymbol{x}_n^k : \forall t\}$, with $\bar{\boldsymbol{\mu}}_{\boldsymbol{X}^k} \subset \mathbb{R}^{n_S \times 1}$, is the $n-$th stochastic feature vector mean along the time, $\overline{\boldsymbol{\Sigma}}_{\boldsymbol{X}^k}(n,l) = (\boldsymbol{x}_n^{k\top} \boldsymbol{x}_l^k)\delta(n-l)$, $\{n,l\} = 1 \ldots n_S$, with $\overline{\boldsymbol{\Sigma}}_{\boldsymbol{X}^{(m)}} \subset \mathbb{R}^{n_S \times n_S}$, is the diagonal variance matrix of each stochastic feature. In the same way

$$\tilde{\boldsymbol{\mu}}_{\boldsymbol{X}^k} = \boldsymbol{1}_{n_S} \boldsymbol{E}\{\bar{\boldsymbol{\mu}}_{\boldsymbol{X}^k} : \forall n\} \tag{10-6}$$

with $\tilde{\boldsymbol{\mu}}_{\boldsymbol{X}^k} \in \mathbb{R}^{n_S \times 1}$ is the mean value of all stochastic features, and $\widetilde{\boldsymbol{\Sigma}}_{\boldsymbol{X}^k} = \text{diag}\{\overline{\boldsymbol{\Sigma}}_{\boldsymbol{X}^k}\}$ with $\widetilde{\boldsymbol{\Sigma}}_{\boldsymbol{X}^k} \subset \mathbb{R}^{n_S \times n_S}$, with $\widetilde{\boldsymbol{\Sigma}}_{\boldsymbol{X}^k} \in \mathbb{R}^{n_S \times n_S}$ is a matrix that comprises in the main diagonal the variance of each stochastic feature.
Consequently, obtained optimization function can be solved in the form:

$$\boldsymbol{\Xi}^k \boldsymbol{\Psi}^k = \boldsymbol{\nu}^k \widetilde{\boldsymbol{\Sigma}}_{\boldsymbol{X}^k} \boldsymbol{\Psi}^k, \tag{10-7}$$

where $\boldsymbol{\nu}^k \subset \mathbb{R}^{n_S \times n_S}$ is the set of generalized eigenvalues that along with the projecting matrix eigenvectors are further used to separate the stochastic feature set fulfilling the proposed stationarity restrictions.
It must be remarked that the recursive separation process (RSS-SC) can be implemented as an online methodology using either continuous one sample updates or multiple sample updates as a sliding window, in the latter case, the feature signal representation operates using each sample at a time to update the stochastic representations, so that equal number of rank–one updates are required, while the variability and stationary subspace projection only uses the final stochastic matrix of the operating window, which reduces computational cost and maintains the stationary reconstruction.

11. Validation Results

11.1 Numerical results on simulated data

In the case of simulated data, the correlation index of equation (8-1) is also used as performance measure, however, given the recursive nature of the process, $\rho^{(k)}$ is computed at each new sample data. Experiments of proposed filtration task are carried out under three different power relations, i.e. three SNR levels, that is the power of the non–stationary signal is half the power of the stationary, same power and 1.25 times (SNR$= 3, 0, -1dB$).

The initial sample size of the simulated signals is experimentally fixed to 400, that is, these first samples are used to create the initial stationary subspace as well as the initial stochastic decomposition and the lag parameter L is set to 50 samples. Experiments were carried out using different window sizes, i.e., the number of new available samples was varied to check how well the proposed methodology responded to non–stationarities. Performance is measured similarly to the previous chapter, using correlation index and DS value.

Yet, it was observed that within a small fraction of samples (for simulated data this was observed to be around half the lag size) the separation task performed similarly, thus only results for a 25 samples window (epoch) size are shown. Figure **11-1** shows average correlation index for a total of 600 simulated time–series, where it can be seen that it remains constant over the number epochs. Moreover there is a great improvement of the stationary separation task when compared against the static case of previous section for low levels of SNR.

Methodology is likewise compared against modified ASSA separation method, for which the epoch size and lag parameters are set to 25 and 50 samples respectively in order to make results comparable. This comparison is shown in Figure **11-2** for the three levels of SNR, again it is remarkable how the propose RSS-SC methodology outperforms the results achieved by ASSA in low SNR levels (-1 and 0), as in the last value 3 even when ASSA holds a higher average correlation index, RSS-SC average results are still comparable with a reduction in the standard deviation value, which indicates a more stable and robust methodology.

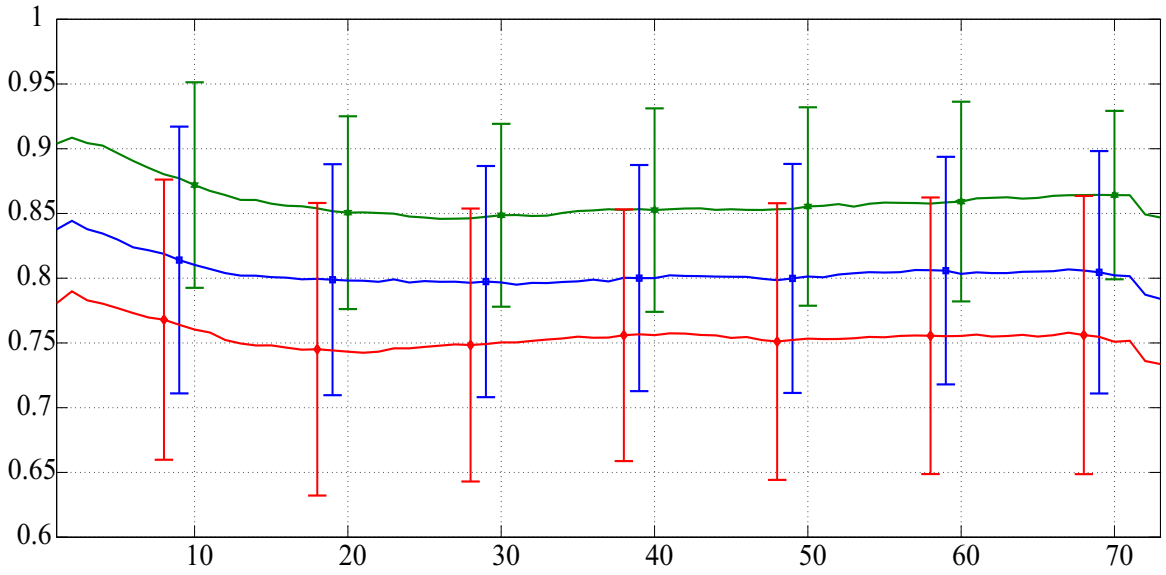


Figure 11-1: Epochs Correlation index evolution for 25 samples step size: (■ SNR = -1dB, ■ SNR = 0dB, and ■ SNR = 3dB)

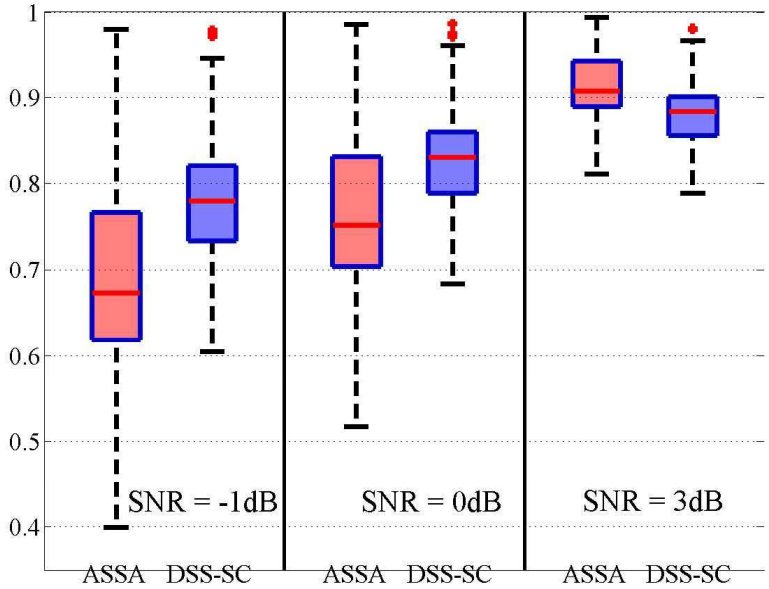


Figure 11-2: Average correlation index for different SNR values.

11.2 Application on EEG Motor Movement/Imagery data

In this case, all nine subjects were used to observed the performance of RSS-SC methodology is separating non–stationary stimuli associated components in EEG data for motor-imagery tasks. On the other hand, the duration of the signal is chosen to be 22.5 seconds, in order to ensure the existence of several stimulus for each type (auditive, visual and motor/imagery). Again Cz channel associated to the sensorimotor cortex is used. For initialization, a 400 samples window size is used with a lag parameter of 60 samples. The epoch size is experimentally set to half the lag size, given the observed results in the synthetic time–series.

Same as before, the degree of stationarity DS is used as performance assessment measure, however this value is computed at each epoch, that is, for a 400 samples window each time new data was collected. Figure **11-3** depicts the signal separation task for a given window as well as the DS value, upper row depicts in black lines the original EEG signal while stationary and non–stationary are depicted in green and red lines respectively. Middle rows show the time-frequency representation of stationary and non–stationary components while last row shows the DS values for all three components.

In this case, given the time-variant nature of the process, a more wide range of frequency is used (0 to $90Hz$), and it is also visible the increment of DS value for the non–stationary component around the $8Hz$ band. This indicates that it is possible then to detect the motor/imagery and stimuli based components on the fly, aiming to automatic annotation BCI systems.

A more illustrative example of this fact its shown in Figure **11-4** where the DS is calculated at each sample epoch and compared against the duration of the trial. First three sub-figures are related with the DS of original signal and separated stationary and non-stationary signals, respectively. As expected, it can be seen higher DS values for the non-stationary signal for each stimulus. At epoch 50, the auditive signal that indicates the beginning of the trial generates high DS values, those values decreases during the rest state of the fixation cross. When the visual cue of the motor/imagery task is given (epoch 67), once again, the DS value increases and falls when the subject has understood which tasks must execute (epoch 77). Finally, during the total duration of the executed task, the DS value holds high values.

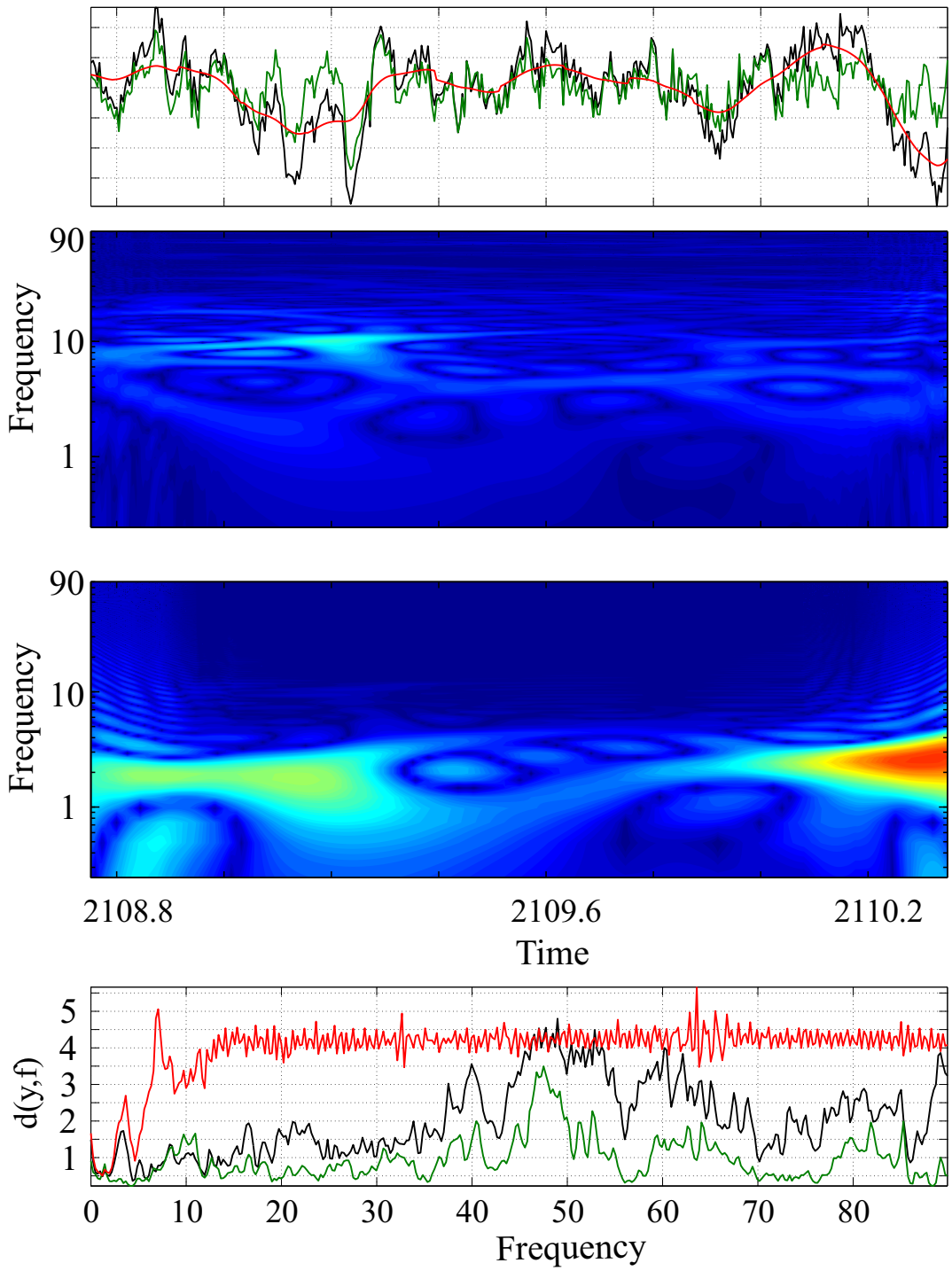


Figure 11-3: Degree of stationarity for motor/imagery epoch, (■ Original, ■ Stationary, and ■ Non-Stationary signal)

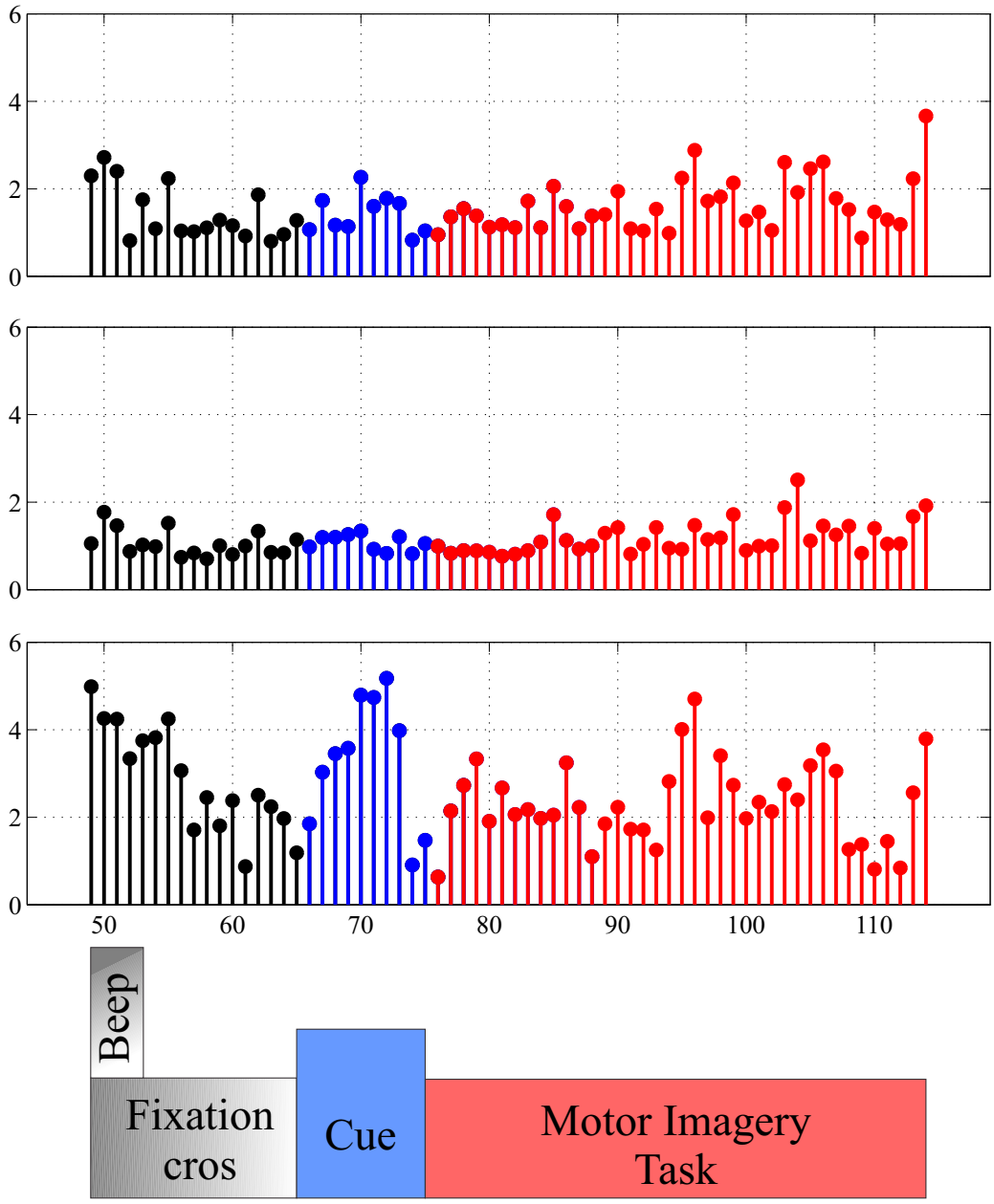


Figure 11-4: Epochs degree of stationarity for motor/imagery trial.

12. Discussion and Conclusions

The main goal of present chapter is to develop a recursive approach for separating between stationary and non-stationary signals from a given time-series. For this purpose, we introduce a recursive Subspace-based Separation with Stationarity Constraints that is based on the hypothesis that once enhancing dynamically the input signal, and projecting the enhanced representation with stationarity constraints, the separation filtering task should be improved. Obtained results allow remarking the following aspects:

i) Introduced recursive time-series enhancement to extract the set of time-variant features of each analyzed window, and it allows highlighting the time-varying statistics of the input signals. In this way, unlike conventional off-line approaches, the recursive decomposition allows identifying different stationary and non-stationary processes that may appear in the analyzed time-series.

ii) Once the decomposition of input segment (epoch) is carried out, the stationarity signal separation appraises the following steps: first, we compute the subspace projection by imposing stationarity constraints that helps to reveal hidden structure behind the characterized time-series as well as to mitigate noise influence.

Finally, we make clear how each projected feature contributes to the stationary of the input time-series. As a result, filtered stationary signal comprises only the amount of each projected feature fulfilling the imposed stationarity constraints.

iii) It is worth noting that we propose a recursive (on–line) approach separating stationary and non-stationary signals. The entire process is carried out by computing an initial decomposition of a small time-series window, and an initial subspace projection along with the respective bias factor of each projected feature, which allows to compute the separation of every epoch. As new data is available, the decomposition is updated via recursive stochastic feature estimation based upon rank-one perturbation analysis, which lowers the needed computational complexity since no new eigen-decomposition in the stochastic components extraction is required. Finally, accuracy rates (correlation index or stationarity degree) are computed over each new epoch.

iv) Proposed approach is tested under different scenarios: i) when we have previous knowledge of the real stationary and non-stationary processes that comprises the input signal (simulated time-series), and ii) when we assume that the input signal has non-stationary behavior in some segments, but we do not have certainty about the nature of the recording (real motor/imagery EEG data). In the former case, we test the proposed approach under several relations between the power of stationary and non-stationary signal (SNR). Additionally, we compared our results against a state of the art approach aimed to solve the same task. Results show that our approach is able to do a better identification of the stationary and non-stationary signals, even for low SNR values. Additionally, we modify the window size length, in order to show sensitivity of our approach to that parameter, nevertheless, we do not find significative differences among obtained results. Hence, we do not show such results.

In the latter case, we use as performance measure the degree of stationarity that is proposed in [39]. In this case, we can see high correspondence between increased non-stationary activity to the presence of different types of stimulus, where peaks of non-stationary behavior are expected. Consequently, the proposed approach through those changes in the stationarity level, could help to infer and identify segments with relevant information or regions related with specific restrictions in recordings with no activity markers (labels).

As the conclusions, it is propose a recursive approach based on Subspace projection with stationary constraints to separate between stationary and non-stationary signals. The method is based on the hypothesis that updating recursively an enhanced representation of the input signal, and mapping those features to a subspace under stationarity constraints, we must be able to obtain a better separation. To assess the validation of our approach, we test it in both simulated and real data. In both cases, obtained results show that the adaptive approach allows identifying several dynamics that may appear in the input data.
As future work, we are seeking to extend the approach for multichannel data that allows us to include separated spaces to solve classification problems.

Part V

Final Remarks

13. Concluding Remarks

13.1 General Conclusions and main Contributions

Design of a cyclo–stationary framework analysis for detecting non–stationarities

In Chapter II the newly introduction of multiple adaptive constraints into the cyclo–stationary analysis for detecting the presence of interferences was proposed. Multiple constraints were derived from external and internal process related parameters, it was demonstrated over the PCG signal analysis that such constraints, improved posterior analysis of segmentations. In particular, adaptive analysis with hard value thresholds prove to be a very efficient methodology for detecting only low quality signals, (low values of SNR), while even when the signals have several noise sources, if the signal to noise ration is high, the methodology can accept the interferences since they do not hold information and do not obscure the cyclo–stationary information, in contrast to the conventional non-constraint approach [24], that has a high rejection rate even for high SNR values.

Separating Non–stationary components from one and multi–dimensional time series

Most approaches for stationary signal separation, require several prior assumptions regarding signal power, stochastic properties or local stationarity. The problem of separating from either one–dimensional or multi–dimensional time series is the treated detailed in Chapter 7, where subspace based representations and multiple stationarity stochastic constraint are used for solving a general separation problem. The optimization functions presented in [21], [5], make strong assumptions over the local stationarity and are only stated for multi-channel time series, which poses a general problem. Additionally, it is a necessary condition to have an approximate information over the number of stationary sources that is expected.
For tackling such problems a modified, more robust assumption of the optimization function is the proposed, using strong diagonal separation, it was possible to extract approximate stationary information, this assumption permitted simultaneously, the generalization of the methodology for one dimensional time-series, using subspace based projections, particularly SSA (given its ability to enhance stationary properties) was used for decompose and analyze

stationary component of input data. The introduction of the variability measured overcame with the number of sources drawback, since it automatically bias the projected components in the stationary subspace. This multiple system of constraints, was tested over simulated and real data, use of one or multiple time-series simultaneously posed no problem for separation of stationary signals under the proposed scheme and in general it outperformed the state of the art methods ASSA [21] and SONS [1] with only a minimum of increase in computational burden.

Posing online and adaptive scheme for non–stationary analysis in cyclo and stationary processes

Finally, Chapter IV briefly extent the stationary separation methodology into a recursive scheme, such that online separation procedure is possible. Introduction of recursive or adaptive subspace representations into a separation or extraction procedure, arises then as online solutions for filtering tasks. Regarding cyclo–stationary analysis of Chapter II it can be extended in the same manner, the adaptive window analysis, only requires an initial window size and would evolve as more cyclic data is available, while the used assessment signatures are checked only in the temporal vicinities within portions of the signals. Moreover, if an online version of either both methodologies is to be implemented into an automatic diagnosis system for digital signal processing, it would serve as main tool for analyzing the acquisition condition in real time, since performance measures regarding the presence of non–stationary components such as DS value could provide red flag alert whenever quality of the recorded signal lowers.

13.2 Future Work

Cyclo–stationary Blind Component Extraction

The general outline for analyzing non–stationary components over cyclic process presented in this work, aimed directly only at detection of such components, however the stochastic and statistic constraints used for cyclo and purely stationary processes are similar in some sense. Thereby, one might pose a modified version of the separation problem using cyclo–stationarity constraint of Definition (4.1.1), yet, state of the art methodologies regarding this problem posed as a primary condition that cyclo–stationary source separation must be approach from a time-frequency domain. Hence a transformation of the separation problem is required as primary condition in this line of analysis.

Multidimensional Online Stationary Component Separation

The general subspace based representations used through this thesis were tested either under single observation (one channel) time-series or multiple observations (of one channel). Yet, several multichannel decomposition methodologies exist, and use internal multidimensional analysis for unfolding hidden stochastic properties of the input data, such a tensor analysis and multichannel SSA. The introduction of such decomposition techniques into the proposed separation methodology is then the next step regarding this topic, such generalization would serve as tool for several real world applications, such as inverse problem solutions in brain imaging or fault diagnosis and localization.

Generalized outline for measuring degree of stationarity, non–stationarity and cyclo–stationarity processes

There is an increasing need of performance measures that are comparable between them, for example DS value uses specific time-frequency representations, and its not comparable with sliding second and higher order statistics, such as sliding variance, kurtosis or skewness, even though they both are measuring the temporal variation of the data. It is necessary then to provide with a general framework that transforms this measures into a common domain, for example cyclo–stationary analysis commonly uses spectral kurtosis for separating and evaluating the spectral autocoherence of cyclic process. However, in spite of being in the same domain (time and frequency) DS and spectral kurtosis are yet not quite comparable, since there is not a clear form of normalization.

Part VI

Appendixes

A. Variability measure of linearly projected stochastic feature set

We will consider that the stochastic reduction dimension of the projected stochastic feature set can be carried out as a linear combination of $q<p=\text{rank}\{\boldsymbol{X}_{\boldsymbol{\Psi}}\}$ independent basis functions, where the minimum mean squared error (MSE) is assumed as the evaluation measure of reduced representation. To this end, we note each column vector of matrix $\boldsymbol{X}_{\boldsymbol{\Psi}}$ as $\boldsymbol{\chi}_n(t) \in \mathbb{R}^{n_O \times 1}$ the t-th time instant of the n-th stochastic feature, measured over all the observations. Hence, an orthogonal vector set is estimated whose resultant of q weighted linear combination can approximate $\boldsymbol{\chi}_n(t)$, being q as small as possible, in such a way that data information is maximally preserved, that is,

$$\min ||\boldsymbol{\chi}_n(t) - \hat{\boldsymbol{\chi}}_n(t)||^2, \tag{A-1}$$

where $\hat{\boldsymbol{\chi}}_n(t)$ is the reconstruction of $\boldsymbol{\chi}_n(t)$ and $||\cdot||^2$ is the norm squared value.
We can certainly assume that the following orthogonal approximation holds:

$$\boldsymbol{\chi}_n(t) = \sum_{i=1}^{p} c_i(t)\boldsymbol{v}_i, \tag{A-2a}$$

$$\hat{\boldsymbol{\chi}}_n(t) = \sum_{i=1}^{q} c_i(t)\boldsymbol{v}_i, \tag{A-2b}$$

where $\{c_i(t) \in \mathbb{R}\}$ is the representation coefficient set, at the time moment t, and $\{\boldsymbol{v}_i \in \mathbb{R}^{n_O \times 1}\}$ is the set of orthogonal basis functions.
Consequently, taking into account Eqs. (A-2a) and (A-2b), the MSE value optimizing (A-1) can be estimated as:

$$\min(\boldsymbol{\chi}_n(t) - \hat{\boldsymbol{\chi}}_n(t))^\top(\boldsymbol{\chi}_n(t) - \hat{\boldsymbol{\chi}}_n(t)) = \min\left\{\sum_{i=q+1}^{p}(c_i(t)\boldsymbol{v}_i)^\top \sum_{i=q+1}^{p} c_i(t)\boldsymbol{v}_i\right\}, \tag{A-3}$$

Optimization function in (A-3) can be minimized when maximizing its complement. Specifically, the following expression takes place:

$$\max\left\{\left(\sum_{i=1}^{q} c_i(t)\boldsymbol{\upsilon}_i\right)^{\top}\left(\sum_{i=1}^{q} c_i(t)\boldsymbol{\upsilon}_i\right)\right\} = \max\left\{\sum_{i=1}^{q} c_i^2(t)\right\} \tag{A-4}$$

To compute both the representation coefficient and orthogonal basis sets, we perform the following singular value decomposition $\boldsymbol{X}_{\boldsymbol{\Psi}} = \sum_{\forall i} \lambda_i \boldsymbol{\upsilon}_i \boldsymbol{\varsigma}_i$, where $\lambda_i \in \mathbb{R}$ is the i-th singular value, and $\boldsymbol{\upsilon}_i \in \mathbb{R}^{n_O \times 1}$ and $\boldsymbol{\varsigma}_i \in \mathbb{R}^{n_S T \times 1}$ are the i-th left and right singular vectors, respectively. Thus, representation coefficients in Eqs. (A-2a) and (A-2b) are to be derived as $c_i(t) = \lambda_i \varsigma_i(t)$, being $\varsigma_i(t)$ the t-th position of the i-th right eigenvector, at time moment t. As a result, once all the observations are considered, the following variability measure of the n-th stochastic feature at the t-th time instant is obtained from Eq.(A-4):

$$g_n(t) = \boldsymbol{E}\{\lambda_i^2 \varsigma_i(t)^2 : \forall i \in q\} \tag{A-5}$$

Time-variant behavior of variability measure $g_n(t)$ may determine whether the corresponding stochastic feature is stationary or not. In other words, if inertia of a given column-wise feature, $\boldsymbol{\chi}$, keeps low throughout time axis, then, one may assume its stationary nature, otherwise not.

B. Academic Discussion

Within the development of the present thesis the following academic products were obtained:

2014:

[1]. Juan David Martinez Vargas, Cristian Castro Hoyos, Germán Castellanos. *"Recursive separation of stationary components by subspace projection and stochastic constraints."*. Accepted: 22nd International Conference on Pattern Recognition, ICPR 2014.

2013:

[2]. Cristian Castro Hoyos, Santiago Murillo Rendon, Germán Castellanos. *"Heart Sound Segmentation in Noisy Environments"*. Published In: 5th International Work-Conference on the Interplay Between Natural and Artificial Computation, IWINAC 2013, Mallorca, Spain, June 10-14, 2013.

[3]. Santiago Murillo Rendon, Cristian Castro Hoyos, Carlos Travieso Gonzales, Germán Castellanos. *"Phonocardiography Signal Segmentation for Telemedicine Environments"*. Published In: 12th International Work-Conference on Artificial Neural Networks, IWANN 2013, Puerto de la Cruz, Tenerife, Spain, June 12-14, 2013

[4]. Andres Eduardo Castro Ospina, Cristian Castro Hoyos, Diego Peluffo Ordoñez, Germán Castellanos. *"Novel heuristic search for ventricular arrhythmia detection using normalized cut clustering"*. Published In: Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE, Osaka, Japan, 3-7 July, 2013.

2012:

[5]. Cristian Castro Hoyos, Diego Peluffo Ordoñez, Germán Castellanos. *"Constrained affinity matrix for spectral clustering: A basic semi-supervised extension"*. Published In: XVII SYMPOSIUM OF IMAGE, SIGNAL PROCESSING AND ARTIFICIAL VISION - STSIVA 2012, Medellín, Colombia.

2011:

[6]. Cristian Castro Hoyos, Diego Peluffo Ordoñez, Germán Castellanos. *"On the groups number estimation for unsupervised clustering"*. Published In: XVI SYMPOSIUM OF IMAGE, SIGNAL PROCESSING AND ARTIFICIAL VISION - STSIVA 2011, Medellín, Colombia.

[7]. Francisco Javier Martínez, Cristian Castro Hoyos, Diego Peluffo Ordoñez, Germán Castellanos. *"Identification of cardiac arrhythmias by means of wavelet packet-based features,"*. Published In: 7th International Seminar on Medical Information Processing and Analysis SIPAIM 2011, 5-7 December 2011, Universidad Industrial de Santander, Bucaramanga, Colombia

Bibliography

[1] Abdi, Hervé ; Williams, Lynne J.: Principal component analysis. In: *Wiley Interdisciplinary Reviews: Computational Statistics* 2 (2010), Nr. 4, S. 433–459

[2] Ahlstrom, Christer: *Nonlinear Phonocardiographic Signal Processing*, Institute of Technology, Linköping University, Dissertation, 2008

[3] Antoni, Jérôme: Cyclostationarity by examples. In: *Mechanical Systems and Signal Processing* 23 (2009), Mai, Nr. 4, S. 987–1036. – ISSN 08883270

[4] Araki, Shoko ; Nesta, Francesco ; Vincent, Emmanuel ; Koldovský, Zbyněk ; Nolte, Guido ; Ziehe, Andreas ; Benichoux, Alexis: The 2011 Signal Separation Evaluation Campaign (SiSEC2011): - Audio Source Separation -. In: *Latent Variable Analysis and Signal Separation* Bd. 7191, Springer Berlin Heidelberg, 2012. – ISBN 978–3–642–28550–9, S. 414–422

[5] von Bünau, Paul ; Meinecke, Frank C. ; Király, Franz C. ; Müller, Klaus-Robert: Finding Stationary Subspaces in Multivariate Time Series. In: *Phys. Rev. Lett.* 103 (2009), Nov, S. 214101

[6] Castro Hoyos, C. ; Murillo-Rendón, S. ; Castellanos-Dominguez, C.G.: Heart Sound Segmentation in Noisy Environments. In: *Natural and Artificial Models in Computation and Biology* Bd. 7930, Springer Berlin Heidelberg, 2013. – ISBN 978–3–642–38636–7, S. 254–263

[7] Chien, Jen-Tzung ; Hsieh, Hsin-Lung: Nonstationary Source Separation Using Sequential and Variational Bayesian Learning. In: *Neural Networks and Learning Systems, IEEE Transactions on* 24 (2013), Nr. 5, S. 681–694

[8] Choi, Samjin ; Jiang, Zhongwei: Comparison of envelope extraction algorithms for cardiac sound signal segmentation. In: *Expert Syst. Appl.* 34 (2008), Februar, Nr. 2, S. 1056–1069. – ISSN 0957–4174

[9] Cichocki, Andrzej ; ichi Amari, Shun: *Adaptive Blind Signal and Image Processing*. 1. Wiley, 2002

[10] Cioch, W. ; Knapik, O. ; Leśkow, J.: Finding a frequency signature for a cyclostationary signal with applications to wheel bearing diagnostics. In: *Mechanical Systems and Signal Processing* 38 (2013), Juli, Nr. 1, S. 55–64. – ISSN 08883270

[11] Delgado-Trejos, Edilson ; Quiceno-Manrique, A. F. ; Godino-Llorente, J. I. ; Blanco-Velasco, M. ; Castellanos-Dominguez, G.: Digital Auscultation Analysis for Heart Murmur Detection. In: *Annals of Biomedical Engineering* (2009)

[12] Gardner, W.A.: *Introduction to random processes: with applications to signals and systems*. Second. Macmillan Pub. Co., 1986. – ISBN 9780029487907

[13] Gardner, W.A.: *Cyclostationarity in communications and signal processing*. illustrated. IEEE Press, 1994 (Electrical engineering, communications and signal processing). – ISBN 9780780304345

[14] Gardner, William A. ; Napolitano, Antonio ; Paura, Luigi: Cyclostationarity: Half a Century of Research. In: *Signal Process.* 86 (2006), April, Nr. 4, S. 639–697. – ISSN 0165–1684

[15] Golyandina, N. ; Nekrutkin, V. ; Zhigljavsky, A.A.: *Analysis of Time Series Structure: SSA and Related Techniques*. illustrated. Taylor & Francis, 2001 (Chapman & Hall/CRC Monographs on Statistics & Applied Probability). – ISBN 9781420035841

[16] Golyandina, Nina ; Zhigljavsky, Anatoly: *Singular Spectrum Analysis for Time Series*. 2013. Berlin, Heidelberg : Springer Berlin Heidelberg, 2013

[17] Gomez-Garcia, J. ; Godino-Llorente, J. ; Castellanos-Dominguez, G.: Non-uniform Embedding based on Relevance Analysis with Reduced Computational Complexity. Application to the detection of pathologies from biosignal recordings. In: *Neurocomputing,* (Accepted)

[18] Gouy-Pailler, C. ; Congedo, M. ; Brunner, C. ; Jutten, C. ; Pfurtscheller, G.: Nonstationary Brain Source Separation for Multiclass Motor Imagery. In: *Biomedical Engineering, IEEE Transactions on* 57 (2010), Nr. 2, S. 469–478

[19] Gu, Fanglin ; Zhang, Hang ; Zhu, Desheng: Blind separation of non-stationary sources using continuous density hidden Markov models. In: *Digit. Signal Process.* 23 (2013), September, Nr. 5, S. 1549–1564. – ISSN 1051–2004

[20] Hara, Satoshi ; Kawahara, Yoshinobu ; Washio, Takashi ; von Bünau, Paul ; Tokunaga, Terumasa ; Yumoto, Kiyohumi: Separation of stationary and non-stationary sources with a generalized eigenvalue problem. In: *Neural networks : the official journal of the International Neural Network Society* 33 (2012), September, S. 7–20. – ISSN 1879–2782

[21] Hara, Satoshi ; Kawahara, Yoshinobu ; Washio, Takashi ; von Bünau, Paul ; Tokunaga, Terumasa ; Yumoto, Kiyohumi: Separation of stationary and non-stationary sources with a generalized eigenvalue problem. In: *Neural Networks* 33 (2012), Nr. 0, S. 7 – 20

[22] Hopgood, J.R. ; Rayner, P. J W.: Single channel nonstationary stochastic signal separation using linear time-varying filters. In: *Signal Processing, IEEE Transactions on* 51 (2003), Nr. 7, S. 1739–1752

[23] Kumar, D. ; Carvalho, P. ; Antunes, M. ; Henriques, J. ; Sa e Melo, A. ; Schmidt, R. ; Habetha, J.: Third Heart Sound Detection Using Wavelet Transform-Simplicity Filter. In: *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*, 2007. – ISSN 1557–170X, S. 1277 –1281

[24] Kumar, D ; Carvalho, P ; Antunes, M ; Paiva, R P. ; Henriques, J: Noise detection during heart sound recording using periodicity signatures. In: *Physiological Measurement* 32 (2011), Nr. 5, S. 599

[25] Luo, Yuhui ; Wang, Wenwu ; Chambers, J.A. ; Lambotharan, S. ; Proudler, I.: Exploitation of source nonstationarity in underdetermined blind source separation with advanced clustering techniques. In: *Signal Processing, IEEE Transactions on* 54 (2006), Nr. 6, S. 2198–2212

[26] Martínez-Vargas, J. ; Sepulveda-Cano, L. ; Travieso-Gonzalez, C. ; Castellanos-Dominguez, G.: Detection of obstructive sleep apnoea using dynamic filter-banked features. In: *Expert Systems with Applications* 39 (2012), Nr. 10, S. 9118 – 9128. – ISSN 0957–4174

[27] Mirmomeni, M. ; Lucas, C. ; Araabi, B.N. ; Moshiri, B. ; Bidar, M.R.: Recursive spectral analysis of natural time series based on eigenvector matrix perturbation for online applications. In: *Signal Processing, IET* 5 (2011), Nr. 6, S. 515–526. – ISSN 1751–9675

[28] Panjer, H.H. ; Society, A.M.: *Actuarial Mathematics*. illustrated. American Mathematical Society, 1986 (AMS short course lecture notes). – ISBN 9780821867426

[29] Pfurtscheller, G. ; Brunner, C. ; Schlogl, A. ; da Silva, F.H. L.: Mu rhythm (de)synchronization and EEG single-trial classification of different motor imagery tasks. In: *NeuroImage* 31 (2006), Nr. 1, S. 153 – 159. – ISSN 1053–8119

[30] Schillinger, Dominik ; Stefanov, Dimitar ; Stavrev, Atanas: The method of separation for evolutionary spectral density estimation of multi-variate and multi-dimensional non-stationary stochastic processes. In: *Probabilistic Engineering Mechanics* 33 (2013), Nr. 0, S. 58 – 78

[31] Sepulveda-Cano, LinaMaria ; Acosta-Medina, CarlosDaniel ; Castellanos-Dominguez, Germán: Finite Rank Series Modeling for Discrimination of Non-stationary Signals. In: *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications* Bd. 7441. Springer Berlin Heidelberg, 2012. – ISBN 978–3–642–33274–6, S. 691–698

[32] Sörnmo, L. ; Laguna, P.: *Bioelectrical Signal Processing in Cardiac and Neurological Applications*. revised. Academic Press, 2005 (Academic Press Series in Biomedical Engineering). – ISBN 9780124375529

[33] Stankovic, L. ; Orovic, I. ; Stankovic, S. ; Amin, M.: Compressive Sensing Based Separation of Nonstationary and Stationary Signals Overlapping in Time-Frequency. In: *Signal Processing, IEEE Transactions on* 61 (2013), Nr. 18, S. 4562–4572

[34] Suresh, P. ; Venkataramaniah, K. ; Thayaparan, T.: Separation of nonstationary signals using Fourier Bessel, fractional Fourier and time-frequency analysis. In: *Signal Processing Image Processing Pattern Recognition (ICSIPR), 2013 International Conference on*, 2013, S. 82–86

[35] Tang, Hong ; Li, Ting ; Qiu, Tianshuang: Cardiac cycle detection for heart sound signal based on instantaneous cycle frequency. In: *2011 4th International Conference on Biomedical Engineering and Informatics (BMEI)* 2 (2011), Oktober, Nr. 3, S. 676–679. ISBN 978–1–4244–9352–4

[36] Tang, Hong ; Li, Ting ; Qiu, Tianshuang ; Park, Yongwan: Segmentation of heart sounds based on dynamic clustering. In: *Biomedical Signal Processing and Control* 7 (2012), September, Nr. 5, S. 509–516. – ISSN 17468094

[37] Tangermann, Michael ; Müller, Klaus-Robert ; Aertsen, Ad ; Birbaumer, Niels ; Braun, Christoph ; Brunner, Clemens ; Leeb, Robert ; Mehring, Carsten ; Miller, Kai J. ; Mueller-Putz, Gernot ; Nolte, Guido ; Pfurtscheller, Gert ; Preissl, Hubert ; Schalk, Gerwin ; Schlögl, Alois ; Vidaurre, Carmen ; Waldert, Stephan ; Blankertz, Benjamin: Review of the BCI Competition IV. In: *Frontiers in Neuroscience* 6 (2012), Nr. 55. – ISSN 1662–453X

[38] Tokunaga, Terumasa ; Kohta, Hiroko ; Yoshikawa, Akimasa ; Uozumi, Teiji ; Yumoto, Kiyohumi: Global features of Pi 2 pulsations obtained by independent component analysis. In: *Geophysical Research Letters* 34 (2013), Nr. L14106, S. 1–6

[39] Tong, S. ; Li, Z. ; Zhu, Y. ; Thakor, N. V.: Describing the Nonstationarity Level of Neurological Signals Based on Quantifications of Time-Frequency Representation. In: *IEEE Transactions of Biomedical Engineering* 54 (2007), Nr. 10, S. 1780–1785

[40] Urbanek, Jacek ; Barszcz, Tomasz ; Antoni, Jerome: Time–frequency approach to extraction of selected second-order cyclostationary vibration components for varying operational conditions. In: *Measurement* 46 (2013), Mai, Nr. 4, S. 1454–1463. – ISSN 02632241

[41] Vargas, Juan David M.: *Estimación de características en superficies tiempo-frecuencia orientada a la detección de patologías en bioseñales*, Universidad Nacional de Colombia - Sede Manizales, Dissertation, 2011

[42] Vitanov, Nikolay K. ; Sakai, Kenshi ; Dimitrova, Zlatinka I.: SSA, PCA, TDPSC, ACFA: Useful combination of methods for analysis of short and nonstationary time series. In: *Chaos, Solitons & Fractals* 37 (2008), Nr. 1, S. 187 – 202. – ISSN 0960–0779