UNIVERSITA' DEGLI STUDI DI PISA

Facoltà di Ingegneria

Corso di laurea in Ingegneria Biomedica

Tesi di Laurea Specialistica

QUATERNION-BASED COMPLEXITY STUDY OF HUMAN POSTURAL SWAY TIME SERIES

Candidato: Andrea Bravi

Relatori: Angelo Maria Sabatini

Luigi Landini

ANNO ACCADEMICO 2008 – 2009

Ognuno di voi ha contribuito alla realizzazione di questo lavoro. In realtà è un po' riduttivo… ognuno di voi ha contribuito alla mia crescita personale e professionale. Ognuno a modo suo. Standomi vicino, prendendomi in giro, facendomi arrabbiare, rendendomi felice o triste, con un sorriso o un insulto…
Si fa per dire! ☺

Dedico quindi questo lavoro a tutti voi:
Alla mia famiglia (parenti inclusi!), ai miei amici e a Giulia… se non ci fosse lei!!!

Un piccolo traguardo altrimenti irraggiungibile

Abstract

Many of the components involved in postural control are not considered while studying the body sway, because classified as unchangeable. Moreover, the anteroposterior (AP) and mediolateral (ML) time series are usually studied separately. In this work of thesis we neglected those reductionist approaches, investigating two main features of the Center of Pressure (CoP):

- The effect of the cognitive component of the subject on the postural sway
- The dimensionality of the center of pressure, through a complex- and quaternion-based representation

The knowledge of the effects of the mind activity on the sway is mandatory to be able to discriminate between mal engineered and proper records. In fact, the attention of the subject on his postural control modifies the features of the CoP signal, conditioning the results of the study. This effect was evaluated recurring to a time-fixed and a time-dependent study of several parameters. The calculations included the Hurst exponent, the Fuzzy Entropy, the Delay Vector Variance and the Recurrence Quantification Analysis. The problem of the parameter setting was deeply evaluated mixing those methodologies in a new way, and our case study in the comparison Open Eyes vs Closed Eyes was reported.

The dimensionality problem arises to obtain a deeper understanding of the structure of the postural fluctuations. Indeed, in many case a multidimensional representation of the data achieves better results than a low-dimensional multivariate representation. We evaluated the complex nature of the AP and ML components. The complex representation allows the use of the information taken from one component to improve the understanding of the other. Because of the possibility that a part of the dynamical structure of the system is shared in both the open-eyes and closed-eyes conditions, we analysed also a quaternion-based representation, joining AP and ML time series recorded in both the visual conditions. To compare the two mathematical models, the prediction gain of a M-step ahead predictor was used. The adaptive filters involved in the processing were the Complex LMS, the Quaternion LMS, and their augmented versions ACLMS and AQLMS, designed to take in account the non-circularity of the signal.

Contents

Index of Pictures

Index of Tables

1. Introduction

The upright stance of the human species is an ode to instability: without a balancing process, the body would fall owing to the force of gravity. The study of human postural control aim to improve the understanding of human balance and the evaluation and rehabilitation of individuals with balance disorders. The regulation of balance is an activity important for everyday life, constituted by a complex multisensory feedback process, involving vestibular system, vision, and somatosensation. Many works are focused on the creation of models of the postural control (1), (2), (3).

If the control process gets altered or impaired (by disease or natural degeneration due to aging, for instance) it can have a deleterious impact on balance. In (4) and (5) is presented a comparison between young and elderly reaction to mechanical stimuli during quiet stance. In (6) features characterizing patients recovering from stroke are investigated. Differences between parkinsonians and healthy subjects are outlined in (7), (8). In (9) vestibular impaired people are confronted with healthy subjects.

Together with the physiological influences, it was showed that the consciousness of postural position can somehow modify the control procedure of the body, (6) and (10).

The integrity of the postural control system is typically evaluated in two ways: (3)

- static posturography, stabilogram and statokinesigram
- dynamic posturography

Dynamic posturography characterizes the performance of the postural control system by measuring the postural response to an applied or volitional postural perturbation, while static posturography characterizes the performance of the postural control system in a static position and environment during quiet standing.

Usually, two variables are used to study those posturographies:

- Center of Mass (CoM), the centroid of the mass elements (corporal segments) that constitute the body. It reflects the actual movements of those segments.
- Center of Pressure (CoP), the centroid of the pressions applied by the feet surface that is in contact with the base of support. It is the point of application of the resulting force exchanged between the feet and the ground, reflecting the action of the muscular forces. The position of CoP determines the moment arm of the reaction force (Figure 1).

Biomechanics studies discovered two types of mechanical control:

- On the CoP, acting through the activation of the plantaflexor muscles and the dorsal flexor of the ankle. That is called ankle strategy, and it is usually used for small perturbations.
- On the CoM, through the movement of the corporal segments, in particular of the trunk. That is called hip strategy, it is usually used for large perturbations.

The CoP is analysed more often of the CoM, because the estimation of the latter one by posturographic data it is inaccurate. The accuracy depends on the knowledge of the anthropometric and inertial parameters, on the error propagation of the measures and the validity of the hypothesis of the applied model.

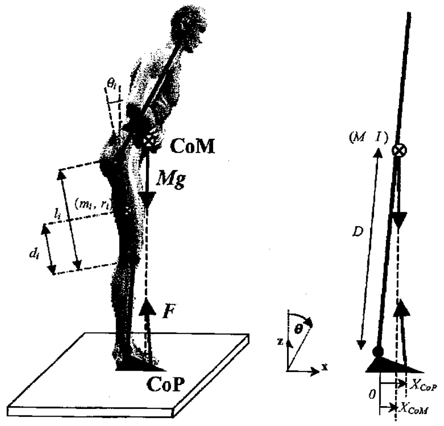


Figure 1 - CoP and CoM representations on a human body (3)

To reduce the computational effort and simplify the analyses, usually the CoP is studied separating the signal in its components anteroposterior (AP) and mediolateral (ML). The bidimensional representation of the CoP is called statokinesigram, while the monodimensional representation of its components is called stabilogram (Figure 2).

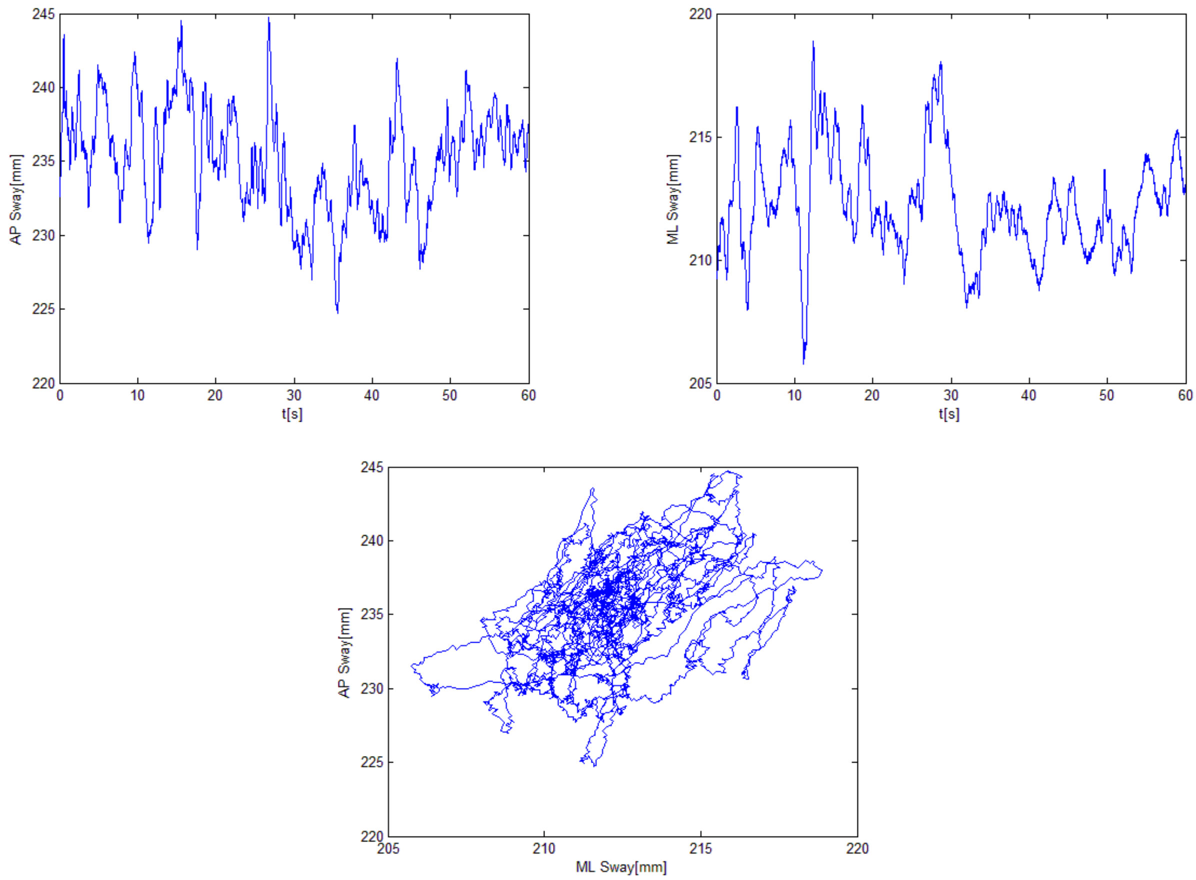


Figure 2 - Example of anteroposterior and mediolateral stabilograms (top), together with the relative statokinesigram (bottom)

Postural steadiness evaluations often include different experimental conditions, stimulating the senses or the mind activity of the examined subject. Usually, eyes-open and eyes-closed trials are used to estimate the role of the visual system in maintaining standing balance, (6),(10), (11), (12). The ratio of the eyes-closed measure to the eyes-open measure is referred to as the Romberg ratio. In some cases, haptic support is given to the patients, or the natural position for standing is altered to see the different behaviors in the control process (13). In most recent studies, also the effect of the mental activity on the CoP was studied (10).

As reported in (3), it is possible to create a block diagram of the postural control to better understand the effect of different perturbations or impairment:

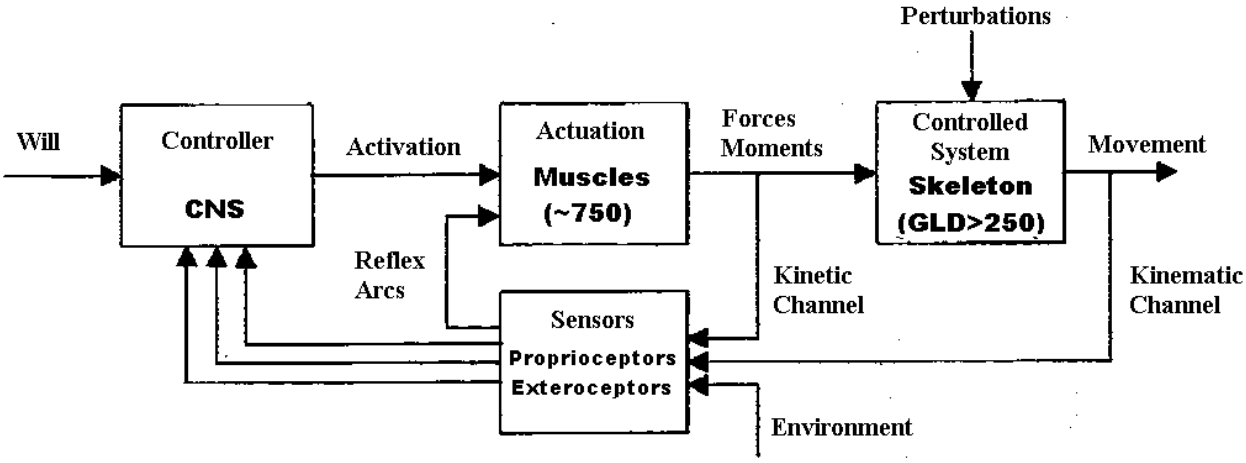


Figure 3 - Control scheme of postural movements

Thus, more into detail, the components that must be considered are:

- Sensorial system: vestibular, somatosensory and visual.
- Orientation perception: Midbrain, Thalamus, Paretal cortex
- Predictive skill of the central nervous system (CNS)
- Cognitive component of the CNS
- Motor coordination
- Biomechanics of the muscular-skeleton system: forces production and intensity scaling
- Adaptation to the environment

Usually, the vestibular system, together with the orientation perception, the predictive skills and the cognitive component of the CNS, the motor coordination and the biomechanics of the muscular-skeleton systems are considered to be unchanged during the posturography. Somatosensory and visual systems are the only ones that are considered time-dependent, depending on the specific protocol applied in the study. Thus, the environmental changes are to be considered as the external stimuli that can be applied to the subject.

In this work of thesis, we tried not to recur to the reductionist approach presented above, widely used in the literature. We decided to understand the effect of the cognitive component on the posturography, to be able to discriminate between mal engineered and proper records. In fact, as showed in (10), (14) and argued in (6), the attention of the subject on his postural control modifies the features of the CoP signal, conditioning the results of the study.

Together with the effect of the mind activity, we investigated the dimensionality of the CoP signal through the modeling of the AP and ML time series of the CoP in the complex ((15),(16)) and the quaternion domains (17). Again, the purpose was not to reduce the complexity of the signal as done in literature. Indeed, in most of the studies just AP time series are considered. When both AP and ML are elaborated, to the best of our knowledge, they were studied always as two separated signals. Furthermore, records in the open eyes and closed eyes conditions are always compared, but never used together to create an exhaustive model of the postural control. The complex- and quaternion-based representations model an exchange of information between the dimensions, that could be the key to better understand and represent the postural control system (18).

To achieve the presented objectives, we relied on many algorithms taken from the elaboration of Chaotic and Stochastic systems. Thus, we investigated the fractal structure of the CoP. Indeed, fractality is one of the key components, shared by both the typology of systems. Because nowadays it is still difficult to separate the two domains, as depicted in Figure 4, we mixed the methodologies trying to create new ways to solve the problem of the parameters setting, to gain a deeper understanding of the phenomena underlying the postural control.

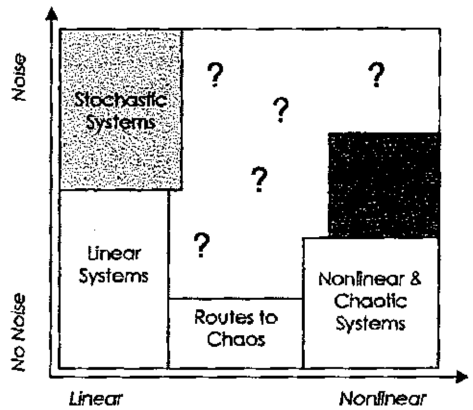


Figure 4 - Nonlinearity and Stocasticity boundaries

2. Theoretical Background

Fractality

An object is said to have fractal properties when it express two features:

- Self-similarity is one of the main features. That means that the fractal object repeats itself many times at different scales (i.e. the system is scale-invariant). For real fractals, the copies of the process at different scales will not be identical, but statistically similar.
- Non-integer (fractal) dimension. The fractal dimension characterizes how the object fills its space. In addition, it describes how the object scales. The scaling of the fractals follows a power law.

Many are the example of fractal systems. In the case of "artificial" systems, they can be mathematical models like the Koch curve, or homemade examples of contemporary art (Figure 5).

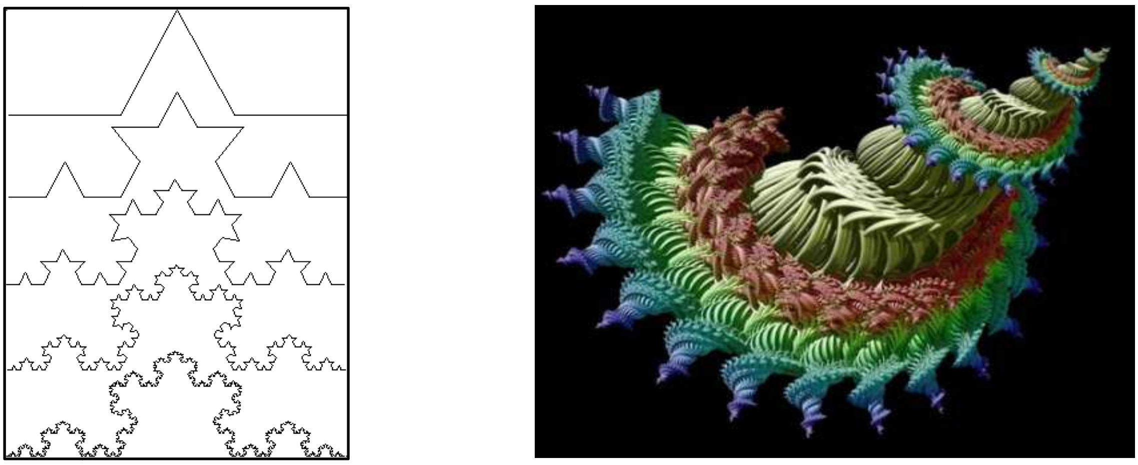


Figure 5 - A Koch curve ad different scales (left) and an example of fractal art (right)

While, in the case of real fractals, many examples at different scales, can be taken from Nature (Figure 6).

Figure 6 - A desert landscape shaped by wind (left), a Roman Cauliflower (middle), a bacterial colony on a Petri dish (right)

Both time and space can be fractal. In fractal time, randomness and determinism, chaos and order coexist. Exactly as in the pictures above, where the global structures are clearly defined but there is a local randomness. An example are electrocardiographic (ECG) data, where the QRS complex structure is known, but there is a variability in the time intervals between two R peaks (Figure 7).

Figure 7 - Detail of a QRS complex (left) and example of ECG record (right)

The Hurst exponent is a measure invented by Hurst to quantify the scaling law of the fractal. More into detail, H is the scaling exponent, and it can assume any real value in the range: $0 < H < 1$. Referred to time series, the value of H can define three important properties:

- If $0.5 < H < 1$, the time series is persistent. Therefore, it is characterized by long memory effects. Being fractal, there is not a characteristic time scale, and theoretically, every sample is correlated with the others.
- If $H = 0.5$, the time series is an independent process, belonging to the family of Brownian motion (the model for a random walk process obtained integrating a white Gaussian noise)
- If $0 < H < 0.5$ the time series is antipersistent. That means that the system covers less distance than a random one, reversing itself more frequently than a random process. Often systems belonging to the antipersistent class present intermittence, the characteristic of a system to periodically pass through stability and instability.

The advantages in having a fractal structure are still a topic of research. However, it is widely accepted the idea that a fractal structure is more stable and less prone to error. The global determinism gives a structure, while the local randomness induces innovation and variety.

The coexistence of both determinism and stochasticity is the most fascinating feature of fractals. Moreover, fractality is one of the features that can be found in both "pure" stochastic and deterministic systems.

Stochastic systems

Taking an arbitrary probability space (Ω, f, P), composed respectively by a set Ω of elementary events (namely ω), a family f of events and a measure of probability P that satisfies the additivity property and

$$0 \leq P(\omega) \leq 1$$

$$P(\Omega) = 1$$

$$P(\emptyset) = 0$$

if a function with real values $X(\omega)$ is measurable, than it is called a random variable. A random variable X is completely described by it distribution function, defined as

$$F(x) = P[X(\omega) \leq x]$$

where x is an arbitrary value. Let's suppose that $F(x)$ is continuous and derivable over all x, it is possible to define the probability density function as

$$p(x) \triangleq \frac{dF(x)}{dx}$$

The Fourier transform of $p(x)$ is called characteristic function:

$$\phi(u) \triangleq \mathrm{E}\{e^{jux}\, dF(x)\} = \int p(x)\, e^{jux}\, dx$$

where $\mathrm{E}\{\cdot\}$ is the expectation operator.

$\phi(u)$ is a complex function of the real variable u. The characteristic function completely defines the random variable X, and from the characteristic function it is possible to calculate the moments of the random variable.

Among the various probability distribution of a random variable, only three are analytically expressible: the normal, the Cauchy and the Lévy distributions. All of them are stable. A random variable is called stable when a linear combination of two independent copies of that variable have the same distribution function of the original, a part for two parameters (see below). The first two are a special case of the latter. For these reasons the family of the Lévy distributions is really important in the analysis of stochastic systems. As a stable distribution, the Lévy characteristic function is defined by four parameters:

- α, the characteristic exponent
- β, the skewness
- c, the scale parameter
- δ, location parameter

Being stable, α and β are preserved, while c and δ can be modified. The first two parameters are the most important, because describe the shape of the probability distribution. α determines the peakedness in δ (the standard normal mean value) and the fatness of the tails of the distribution. β determines the asymmetry. c determines the width of the distribution, it is a measure of dispersion.

When $\alpha = 2$, the Levy distribution collapse into a normal distribution with variance $\sigma^2 = 2c^2$ and mean value δ (β has no effect).

The Lévy distribution is often used to model fractal systems. Indeed, it is called a fractal distribution because of the scale parameter c, which varying creates self-similar copies of the distribution. Using the point of view of fractal system, the characteristic exponent α, which can take fractional values, represents the fractional dimension of the probability space.

Let's consider an example of stochastic fractal processes called $\frac{1}{f}$ noise. That family is characterized by a power spectral density proportional to $\frac{1}{f^b}$, with f the frequence and b the spectral exponent. Depending on the value of the spectral exponent, the time series have different features:

- $b = 0$, white noise. The power spectrum of the time series is not a function of the frequency.
- $b = 2$, brown noise. The noise belongs to the Brownian motion, and the scaling factor is a square.
- $0 < b < 2$, pink noise. The noise is closely related to relaxation processes. A relaxation process is a form of dynamical equilibrium where there is an exchange (loss) of energy between the components of the system. When one rises, the other lowers, and vice versa. An example is the turbulence, where there is an exchange between large scale and small scale structures. Systems with that value
- $b > 2$, black noise. The noise is closely related to long-run cyclical systems. An example is the water level of the Nile river. Two are the black noise-related characteristics that Mandelbrot discovered. The Joseph effect, where the long memory causes the formation of trends and cycles. The Noah effect, where at a certain instant abrupt discontinuities arise (he called them catastrophes); these discontinuities cause the frequency distribution of black noise processes to have high peaks at the mean, and fat tails.

There is a last feature shared by both black noise and pink noise: the mirror effect. As brown noise is the integrand of white noise, black noise is the integrand of pink noise.

Depending whether it has time-dependent variance, or it is a stationary process with a constant expected mean value and constant variance over time, the time series belongs respectively to the family of fractional Brownian motion (fBm) or fractional Gaussian noise (fGn). The mirror effect is true also for these two families; indeed, the fBm is the integrand of a fGn. The scaling law valid for a fBm is:

$$\langle \Delta x^2 \rangle \propto \Delta t^{2H}$$

with Δx^2 the squared displacement and Δt the time interval over which the displacement was observed, and H the Hurst exponent.

The velocity distribution of a fractional Brownian motion belongs to the fGn family, and its probability density function is modeled with a Lévy distribution function.

Finally, we can show the link between the spectral exponent, the Hurst exponent and the scaling exponent:

$$b = 2 * H + 1 \qquad \alpha = \frac{1}{H} \qquad \alpha = \frac{b-1}{2}$$

It is important to remember that, even if α and H are related, their meaning is different. The former describes the dimension of the probability space (related to the statistical self-similarity of the process), while the latter the dimension of the time series.

These results join together everything we said about fractal systems and stochastic systems. The final element that unifies fractality and stochasticity is Chaos.

Deterministic Chaos

Chaotic systems are deterministic, nonlinear dynamical systems. The use of the word "chaotic" refers to the fact that observations of the system, usually made through some measurements, have no discernable regularity or order. To understand what a chaotic system is, it is necessary to understand what is the Phase Space.

The Phase Space is a space of representation of the measurements of the system with dimension m, with m the number of parameters that define the chaotic system. The passage from scalar observations to the multivariate Phase Space is allowed by the embedding theorem, developed by Takens and Mañé. Let's consider a vector $\boldsymbol{x}(n)$, that describes the state of a dynamical system at a certain instant n, which evolution is defined by a certain function F, giving

$$F\big(x(n)\big) = x(n+1)$$

The theorem tells us that if we are able to observe a single scalar quantity $h(\cdot)$, of some vector function of the dynamical variables $\boldsymbol{g}\big(\boldsymbol{x}(n)\big)$, then the geometric structure of the multivariate dynamics can be unfolded from this set of scalar measurements $h\Big(\boldsymbol{g}\big(\boldsymbol{x}(n)\big)\Big)$ in a space made out of new vectors with components consisting of $h(\cdot)$ applied to powers of $\boldsymbol{g}\big(\boldsymbol{x}(n)\big)$. These vectors

$$\boldsymbol{y}(n) = \Big[h\big(\boldsymbol{x}(n)\big), h\Big(\boldsymbol{g}^{\boldsymbol{T}_1}\big(\boldsymbol{x}(n)\big)\Big), h\Big(\boldsymbol{g}^{\boldsymbol{T}_2}\big(\boldsymbol{x}(n)\big)\Big), \dots, h\Big(\boldsymbol{g}^{\boldsymbol{T}_{m-1}}\big(\boldsymbol{x}(n)\big)\Big)\Big]$$

define motion in a m-dimensional Euclidean space, where m is called embedding dimension. That space is called the Phase Space. It was proven that under general conditions of smoothness on $h(\cdot)$ and $\boldsymbol{g}(\boldsymbol{x})$, the evolution in time $\boldsymbol{y}(n) \longrightarrow \boldsymbol{y}(n+1)$, follows that of the unknown dynamics $\boldsymbol{x}(n) \longrightarrow \boldsymbol{x}(n+1)$.

To implement Takens' theorem, usually it is imposed that $h(\cdot)$ is a measurement function of the dynamical system, giving

$$h\big(\boldsymbol{x}(n)\big) = s(n)$$

with $s(n)$ the observed variable of the system. For the general function $\boldsymbol{g}(\boldsymbol{x})$ it is usually chosen the operation which takes some initial vector $\boldsymbol{x}$ to that vector one time delay τ later, so that the T_k^{th} power of $\boldsymbol{g}(\boldsymbol{x})$ is

$$\boldsymbol{g}^{T_k}\big(\boldsymbol{x}(n)\big) = x(n + \tau\, T_k)$$

The components of $\boldsymbol{y}(n)$ will take the form

$$\boldsymbol{y}(n) = [s(n), s(n+\tau), s(n+2\tau), \dots s(n+(m-1)\tau)]$$

Even if strange to think, a single scalar contains enough information to reconstruct a m-dimensional space because in a nonlinear process all the variables are generically connected. Furthermore, the coordinates in the Phase Space hold the information of the time derivates of the signal, when τ is sufficiently small, indeed

$$\frac{ds(t)}{dt} \approx \frac{s(t+\tau) - s(t)}{\tau}$$

$$\frac{d^2s(t)}{dt^2} \approx \frac{s(t+2\tau) - 2s(t+\tau) + s(t)}{\tau}$$

and so on.

Now, having reconstructed a chaotic system in the Phase Space, what we see is that it arises a structure. That is one of the characteristic features of Chaos: structure in the Phase Space. An example is given in Figure 8.

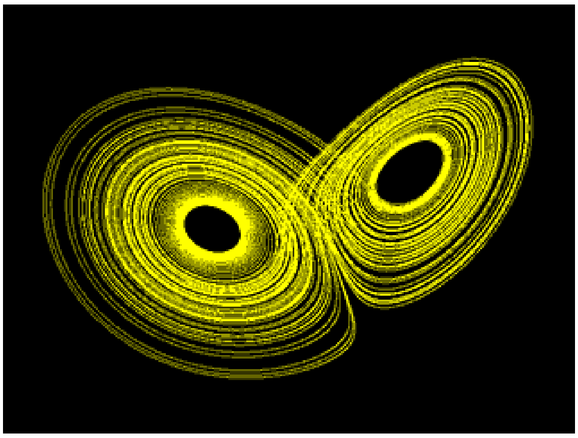


Figure 8 - Lorentz strange attractor

There is depicted the strange attractor of the Lorentz model. The strange attractor is the element that characterizes a chaotic system and that resumes all its properties. Giving definitions, an attractor is a set of points in the Phase Space visited by a dynamical system. A strange attractor is an attractor with fractal dimension. In a strange attractor the points never repeat themselves and the orbits (i.e. a trip around the attractor) never intersect but both the points and the orbits stay within the same region of Phase Space. Strange attractors are characterized by nonperiodic cycles, that means that trips around the attractor has not an absolute frequency but an average frequency. Now we can express what makes a system chaotic. Two are the requirements:

- Existence of a fractal dimension
- Sensitivity to the initial conditions

The strange attractor of a chaotic system shows self-similarity characteristics that fill the phase space according to the fractal dimension. As shown in Figure 8, in the case of the Lorentz model, the self similarity is given by the repeated folding of the attractor. This property links chaos with fractals. The second characteristic results when considering two nearby point or orbits: the two will separate exponentially with a rate of divergence given by an invariant parameter of the Phase Space, called Lyapunov exponent. There is one exponent for each dimension. Chaotic systems have at least one positive Lyapunov exponent. The sum of all the exponents will be negative for dissipative systems, otherwise the system should expand indefinitely in time (i.e. it should have infinite energy). Therefore, even if a chaotic system is deterministic, after a certain time constant given by the Lyapunov exponent the observation of the systems from a certain initial condition becomes incorrelated with the original nearby conditions, creating the illusion to being studying a stochastic system. This property links chaos with stochasticity.

For further insights on chaos, fractals and random systems, refer to (19),(20) and(21).

3. Research statement

We decided to follow the U-model for research (Figure 9) to structure this work of thesis.

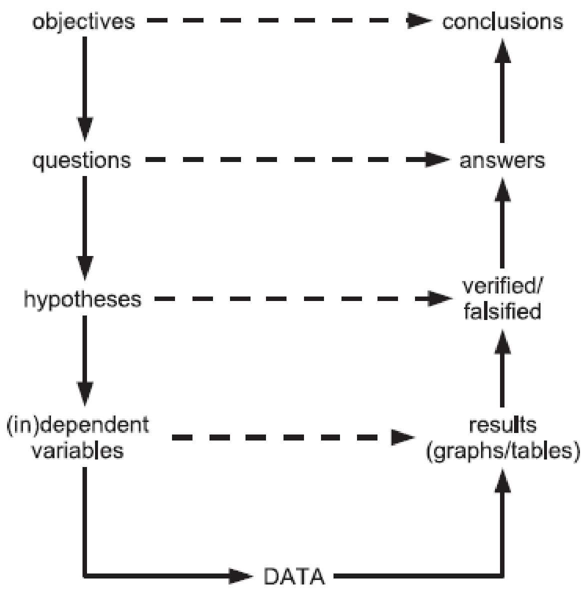


Figure 9 - U-model for research

1. *Objectives*

- Characterization of the CoP time series, studying the degree of non linearity, the degree of determinism, the complexity and the effect of the visual component
- Identification of parameters sensitive to the effects of the cognitive component on the postural sway
- Exploration of the multidimensional nature of the CoP, and investigation of the possible correlations between the closed-eyes and the open-eyes conditions

2. *Research questions*

1) Do anteroposterior and mediolateral time series of the CoP share information, needed to describe in a more complete way the postural control system?

2) Is it possible to associate open-eyes and closed-eyes time series, to be able to model the body sway with a quaternion-based representation?

3) Is the postural control system described in a more complete way relying on the information of both open-eyes and closed-eyes time series, therefore with a quaternion representation, or a complex representation is able to obtain the same information?

3. *Initial hypotheses*

1) The postural control system is composed by the following components:
    - Sensorial system: vestibular, somatosensory and visual.
    - Orientation perception: Midbrain, Thalamus, Paretal cortex
    - Predictive skill of the central nervous system (CNS)
    - Cognitive component of the CNS
    - Motor coordination
    - Biomechanics of the muscular-skeleton system: forces production and intensity scaling
    - Environment interactions

2) During the acquisition process of the CoP (details in 4.1), the effect of all the elements presented above can be classified as constant. The only exception is the cognitive component of the CNS.

4. Methodologies applied

1. *Measure protocol and data acquisition*

In the present work, two typologies of dataset were studied.

The former is the same proposed in (22), with the following protocol. A group of eleven adults (six male, five female; $mean\ age \pm standard\ deviation$ SD: 29.6 ± 4.7, range: $24 - 38$ years, $mean\ weight \pm$ SD: 64 ± 13 kg, range: $50 - 85$ kg; $mean\ height \pm$ SD: 1.72 ± 0.08 m, range: $1.60 - 1.86$ m) was involved in the investigation. All subjects, with no evidence or known history of a gait, postural, or skeletal disorder, provided informed consent prior to participation in the testing protocol. They were instructed to stand with an upright posture on the force platform, with arms relaxed at the side, the feet abducted about 10°. The acquisitions lasted for 60, as suggested by the International Society for Postural and Gait Research (ISPGR). That duration is due to the research of a compromise between the need to obtain the longest time series possible in order to compute a consistent estimation of the dynamical parameters, and the need to obtain the shortest time series possible in order to avoid the tiredness and then the consequent time dependence of the system. Ten trials were conducted on each subject in open-eyes (OE) conditions, with the subjects looking straight ahead at a visual reference; ten trials were conducted with the eyes closed (CE). The order of testing, e.g. OE versus CE trials, was randomised for the subject population. Rest periods of 60 s and 5 min were provided between each trial and between each set of ten trials, respectively. COP data were acquired at 100 Hz, using an AMTi force platform and both the ML component and the AP component were analysed. The first 10 s of each time series were discarded in order to avoid the influence of any transient on data processing (9).

The latter dataset was recorded to test the sensitivity of certain parameters to dynamic behaviors like opening and closing eyes during the acquisition, and focusing and removing the attention (cognitive control) of the subject on postural control. The protocol details will be discussed in the proper section (4.4.4).

The datasets were not filtered mainly for two reasons. As stated in (23), filtered copies of a signal cannot be processed using surrogate data because the null hypothesis of a monotonically rescaled Gaussian linear random process it is usually not true for them. Moreover, in (19), pp. 115-132, it is explained that the effects of linear and non-linear filtering on chaotic data still need to be explored and deeply characterized. Thus, at the moment we are writing, it can be valuable to rely on a well made acquisition platform to reduce noise, avoiding any type of filtering.

2. Time series model

During the processing, those properties of the CoP were considered:

- It is nonstationary
- The velocity of CoP follows a Levy distribution
- The power spectral density is of the $\frac{1}{f^b}$ type, with b the spectral exponent
- The CoP has fractal properties:
  - It is often modeled as a dual-correlated fractional Brownian motion (fBm) random process
  - It has an Hurst exponent (H) in the antipersistent range , thus it is of the pink noise type
- Because of the properties of fBm, integrating the displacement of the CoP with time, the obtained time series has an H in the persistent range, thus it belongs to the black noise type

3. Evaluation of the complex nature of CoP

The first research question presented in chapter 3 was:

Do anteroposterior and mediolateral time series of the CoP share information, needed to describe in a more complete way the postural control system?

To answer, we recurred to the non-parametric test for detecting the complex-valued nature of time series presented in (24). In the article, the authors extend the concept of Delay Vector Variance (25) to the complex domain. Recurring to a Kolmogorov-Smirnoff (K-S test) of the distributions of the DVV plots of the surrogates created through multivariate iterative adjusted amplitude Fourier transform (MViAAFT, (23)) and complex iAAFT (CiAAFT, (24)), it is possible to discriminate the complex nature of the signal.

The complex DVV was applied to the AP and ML time series of the first dataset, confirming the complex-valued nature of the CoP (see paragraph 5.1).

References in the chapter Algorithms:

DVV and Complex DVV see paragraph 8.2
MViAAFT, see paragraph 8.6.2
CiAAFT, see paragraph 8.6.3

4. Sensitivity to the cognitive component

As explained in paragraph 3.3, we made several hypotheses on the blocks constituting the postural control. The only one that we classified as time-dependent was the cognitive component.

Thus, to address the second research question

Is it possible to associate open-eyes and closed-eyes time series, to be able to model the body sway with a quaternion-based representation?

it was needed to understand when the subjects modify the body sway through unconscious movement due to the focusing of their attention on the postural displacement. As proved in (10), (14) and argued in (6), the attention on the postural control can modify the properties of the CoP time series. Obviously, these alterations are related to movements that the subjects do not know to be doing, because during the acquisitions it is explicitly requested them to stand still (we are excluding the hypothesis that the subjects by purpose altered their position).

The idea is that time series with an analogous level of alteration due to cognitive control, can be studied together. In that way, it is possible to associate OE and CE time series.

The study of the effect of the cognitive component is quite new in the literature. Therefore, we mainly relied on (10), paper published in 2007 and (14), published in 2009. In those papers, the most important parameter used to detect the amount of attention was the Sample Entropy (SE). The authors explained that the CE condition has a lower SE of the OE condition, because during the closed-eyes recording the subjects focus more easily their attention on the postural control. Then, when the subjects get distracted with a task during the acquisition, the SE in the CE condition reaches the SE in the OE condition, because the mind is so busy that does not modify anymore the postural sway.

Defined in (26), SE relies on a representation of the data in the Phase Space, with a time delay equal to the sampling time used during the acquisition. Many are the example of application of that methodology; see (27), (28), (29). Anyway, all the cited works neglected to adopt the improvement made in (30), where the embedding reconstruction of the time series for the calculation of the Sample Entropy, was extended to an arbitrary time delay. Recently, as a final improvement in the calculation of Sample Entropy, a new methodology called Fuzzy Entropy (FE) was created (31). FE eliminates the problem of discontinuity due to the use of the Heaviside function, substituting it with a fuzzy membership function, and adds the detrending of the time series to improve the estimation.

In our study, we decided to use FE because of the proven better performances, compared to SE. Furthermore, we added, like in (30), the possibility to choose an arbitrary time delay for the reconstruction on the phase space of the time series.

We decided to conduct two kinds of studies with the purpose to associate properly OE and CE time series:

- Fuzzy Entropy of the whole time series
- Fuzzy Entropy as a function of time

The objective of the first study was to characterize our dataset. In literature many studies presented the evaluation of parameters in the OE and CE conditions; however, the results are often incoherent, even when the chosen protocols and parameters are the same. Usually each author tries to explain his results recurring to physiological argumentation, without justifying the difference with the results of other authors. We decided to create our case and analyse it, to present our point of view.

The objective of the second study was to identify and quantify the amount of attention posed on the postural control, studying windows of the time series. We argued that time series with similar alteration from the mind activity could be studied together (i.e. with a quaternion representation).

References in the chapter Algorithms:

Sample Entropy see 8.3.1
Fuzzy Entropy see 8.3.2

4.1 Parameters choice

Before the usage of FE, it was needed to set the parameters of the algorithm. FE needs of three parameters:

- The time delay for the Phase Space reconstruction, τ
- The embedding dimension for the Phase Space reconstruction, m
- The reference radius for the membership function, r

The time delay was selected recurring to the Average Mutual Information (AMI), as done in (19), (22). Moreover, the assessment of the final values of interest was made using two parameters that we extrapolated from the Delay Vector Variance scatter diagrams (DVV, (25)). One is the maximum deviation from linearity (index of nonlinearity), that is the deviation from the bisector of the scatter diagram; the other is the minimum normalized variance (index of stochasticity), that is the minimum variance explainable by the delay vectors of the original time series.

The embedding dimension was selected recurring to the two parameters taken from DVV, used to calculate the time delay. It was chosen the value that expressed the right compromise between the nonlinearity explanation and the computational load.

The radius was selected using the Recurrence Quantification Analysis (RQA, (32)). That idea is motivated by the similarities between FE and the calculation of the recurrence plots, needed for the estimation of the RQA parameters. Indeed, the calculation of the recurrence plot can be modified to fit exactly the calculation used by the Sample Entropy (SE) and, therefore, the Fuzzy Entropy (FE). The explanation follows. SE and FE rely on a time-delay embedding reconstruction, like RQA. More into the details, the FE and SE are computed averaging the distances between delay vectors (DVs). The metric for the distance is the maximum absolute distance between DVs. RQA is based on the recurrence plot, created calculating the distances between the DVs. If the same metric is used to build the recurrence plots (that usually are based on the Euclidean distance), the result is the recurrence representation of the system studied by SE and FE. Then, the Sample Entropy algorithm uses the Heaviside function to discriminate the DVs distances that are higher or lower than a certain threshold. That is exactly the same mechanism used for the creation of the recurrence plots. Indeed a threshold is applied to classify the point in the recurrence plot as recurrences or non-recurrences. Thus, that threshold is applied to the DVs distances in both the algorithms, and the DVs distances are computed in the same way. Therefore, we are describing exactly the same system with two different methodologies. In the case of FE, the measured distance should be replaced with the relative membership function. Anyway, that would render impossible to estimate the RQA parameters, because FE does not use a threshold (it would be needed to add another threshold, but the meaning of that move would not be clear); it would be just possible to create a recurrence plot with the value of the membership functions in each element of the matrix. We then decided to use the threshold like in the SE, adding the detrending of the time series (like in FE), but avoiding the use of the fuzzy membership function.

Now, the reason why we mixed these methods is clear. RQA allows the quantification of parameters like $\%Determinism$, index of determinism, $\%Recurrence$, index of nonlinear autocorrelation and $\frac{\%Determinism}{\%Recurrence}$, considered to be sensitive to the state alteration of a system. Those parameters are widely used in the postural analysis literature (12), (8), (13) (33). For instance, in (12) the CoPs of the subjects were recorded under OE and CE conditions, together with the head in standard position, and the head facing the side, and $\frac{\%Determinism}{\%Recurrence}$ was able to successfully discriminate the different conditions. We argued that the best threshold in the computing of the Fuzzy Entropy is the one that maximizes the degree of determinism, giving a low value of nonlinear autocorrelation (thus, maximizing entropy). Using this criteria, it can be selected a threshold that tries to minimize the stochastic component of the signal, keeping a high value of entropy.

It is worth to notice that in our case, $\%Recurrence$ is exactly the $B(\varepsilon)$ value used in the calculation of the SE of the detrended time series. The advantage compared with other criteria for the selection of the threshold (an overview is presented in the chapter 8 when SE and FE are introduced) is that this selection process can be extended to any kind of study, because it just relies on the knowledge of the needed characteristics of the time series. In our case, we are orienting the elaboration of the COP toward the characterization of its deterministic features, because it is well known that the stochastic base of the COP is a fractional Brownian motion.

References in the chapter Algorithms:

Average Mutual Information see 8.5.1
Delay Vector Variance see 8.2
Recurrence Quantification Analysis see 8.7
Sample Entropy see 8.3.1
Fuzzy Entropy see 8.3.2

4.2 Fuzzy Entropy of the whole time series

During this phase, the FEs of the time series of the first dataset were computed.

Three are the analyses made:

In the first one, all the subjects Fes were calculated using the time delays, the radii and the embedding dimension selected previously (paragraph 5.5.1). The purpose was to characterize the effect of the visual component (i.e. OE/CE tests) to be able to confront our results with the literature ones. Indeed, many researches published different results on the evaluation of parameters in the OE and CE conditions, and we wanted to present our point of view.

In the second one, we analysed the dependence of the FE to the radius, confronting the values of two specific subjects (paragraph 5.5.2). A comparison between the results obtained during the set of the parameters was made.

In the third one we extended the study to all the time series, discussing the effects of the time delay and the relation between $\%Recurrence$ and FE (paragraph 5.5.3).

4.3 *Fuzzy Entropy as a function of time*

The results of the elaborations of the precedent paragraph were: we argued that the physical differences between subjects justify the differences found in literature. Coherence between the DVV and the RQA parameters was shown, together with an inverse proportionality between FE and $\%Recurrence$. Then, the sensitivity of FE to the radius was outlined, noticing that for certain radii the results obtained can be completely different.

Keeping in mind considerations of the precedent paragraph, we argued that the results of (10) must be analysed differently. The authors explained that the CE condition has a lower SE compared to the OE condition, because during the closed-eyes recording the subjects focus more easily their attention on the postural control. Then, when the subject gets distracted with a task during the acquisition, the SE in the CE condition reaches the SE in the OE condition, because the mind is so busy that does not modify anymore the postural sway. Anyway, because of the inter-variability outlined previously, we think that this result should be considered more generally. Our intuition is that there should be a heightening in the entropy value when a subject goes from the condition of attention to the condition of disregard of the postural control, but that could happen even when the subject has a FE in the CE condition higher than the FE in the OE condition. That was the feature we looked for during the elaborations described in the present paragraph.

The aim was to identify the time series with a considerable amount of distortion due to the attention, to avoid their use, and to find the time series with a low and similar amount of distortion, to combine them in the quaternion model. Two approaches were considered:

- Study of the FE through overlapping windows of the signal
- Study of the sequential Fuzzy Entropy (i.e. the FE at an instant t is calculated considering all the samples before)

In the windowed case, the windows length was chosen remembering that the number of delay vectors used in the FE estimation need to be higher than 50, see (31). Being:

$$N - m * \tau = number\ of\ delay\ vectors$$

with N the number of samples, m the embedding dimension and τ the time delay used to reconstruct the time series in the phase space, for $\tau = 0.2\ s$ and $\tau = 0.5\ s$, the minimum windows length were 150 and 350 samples, respectively (because we used $m = 4$).

In the sequential case, the assumption is that a proper FE calculation needs the history of the process. Again, the values of FE will be acceptable only for time series longer than 150 and 350 samples, as stated before.

The behavior we expected was a bi-stability, for both open-eyes and closed-eyes time series. The assumption was that the “attention-on” and “attention-off” time series had two values of FE more or less fixed. During the elaborations, the correlation between the entropy of the AP and the ML components was pointed. It was assumed sensible to find a coherent effect of the cognitive component on both the time series.

4.4 Study of a specific dataset

The precedent study did not allow a proper classification of the cognitive component, therefore we created a second specific dataset to characterize the dynamical properties of the CoP. The signals were recorded at 140 Hz.

A male adult (25 years, 63 kg, 1.69 m) with no evidence or known history of a gait, postural, or skeletal disorder was involved in the investigation. He provided informed consent prior to participation in the testing protocol. The subject was instructed to stand with an upright posture on the force platform, with arms relaxed at the side, the feet abducted about 10°. Seven time series were recorded under single-blind condition (i.e. the subject did not know the purpose of the experiments):

1. Open-eyes, total concentration of the subject on its body
2. Closed-eyes, total concentration of the subject on its body
3. Alternated open-eyes/closed-eyes conditions
4. Alternated dual-task/single-task, in the CE case
5. Alternated dual-task/single-task, in the OE case
6. Complete dual-task, in the OE case
7. Complete dual-task, in the CE case

Details of tests 1-3:

The subject was asked to totally concentrate on his body, feeling the sensations created by the postural sway. In the third acquisition every 10 seconds the subject was ask to change the visual condition, starting with the OE. The first two time series were recorded for 60 seconds, the last for 120 seconds. The purpose of those acquisitions was to check the sensitivity of the parameters involved in the evaluation of the cognitive component to the changes of the visual component.

Details of tests 4-5:

The dual-task was inspired by (10) and consisted in listening to a word, reversing its letters and uttering back them. Because the speech can modify the CoP, the following procedure was implemented: the subject listened to the word and then waited for 10 seconds; then, 10 seconds were given to him to utter back the reversed word; after those 10 seconds, he had to interrupt is speech or just keep quiet for 20 seconds (single task). The acquisitions started with 10 seconds of single task (that were eliminated during the elaborations to avoid transient, as for the first dataset), then a word was told to the subject and the procedure started. The succession dual-task/single-task was alternated for 120 seconds.

Details of tests 6-7:

The dual-task consisted in listening to the words and reversing them, as before, but this time the subject was asked to remember those words to report them all together at the end of the acquisition (in this way the CoP was not influenced by the speech movements). The acquisitions lasted 120 seconds.

In general:

Words of eight letters were used. The length was chosen by empirically trying to find a right compromise between a difficult and an achievable task. Indeed, it was needed a difficult task to completely deviate the attention of the subject from the postural sway, but an easy one to avoid the subject to get “disappointed”, because his concentration on the task could have been compromised. The sole aim of the presented dual-tasks was to withdraw attention from the postural task. Therefore, the participant was instructed to perform the task to the best of his ability. No feedback on the accuracy with which he performed his task was provided.

The presented dataset was studied with the same time-dependent approach proposed in paragraph 4.4.3. The studied parameters were inspired by (10). Therefore, we first evaluated the use of the embedding parameters found for the precedent dataset, and then we calculated the standard deviation, the sway-path length, the Fuzzy Entropy and the Hurst exponent of the time series. Because of the specificity of the dataset, we studied those parameters also for the differential time series (i.e. velocity of the CoP).

The calculation of the Hurst exponent has been widely discussed in literature, and many are the ways to estimate it. In (10) it was calculated through Detrended Fluctuation Analysis (DFA), while we used the Scaled Windowed Variance (SWV). That choice was due to the results achieved in (34), where a detail procedure explains how to estimate the Hurst exponent. The first step was to calculate the power spectral density (PSD) to assess the limit frequency from which the signal information is mainly due to noise. As a second step, an improved version of the PSD created by the author (the lowPSDwe) is used to classify the signal as a fractional Gaussian noise or a fractional Brownian motion. Then a specific algorithm is used to calculate the Hurst exponent, depending on the family of the signal. Essentially, because the DFA is more suited for fractional Gaussian motion, while the CoP is a signal of the fractional Brownian motion family, we used the SWV (more specifically, we used the linear detrended SWV, an improved version of SWV). A detailed explanation of the procedure is described in 8.0.

5. *Dimensionality of postural sway*

The final purpose of this work was to deepen the understanding of the CoP nature through a dimensionality study.

In literature it can be often found a reductionist approach, where AP and ML time series are studied separately. However the centre of pressure is a bi-variate signal and, as proven during this work, at least a complex representation is needed to create a more accurate model of the postural sway. The complex-based representation introduces statistics that characterize an exchange of information between the dimensions that could represent the key to better model the postural control system (18). Another possibility to create an even more accurate model of postural sway could be to fuse the information taken from both the open-eyes and the closed-eyes condition, recurring to a four dimensional variable.

The components of a biological system or process usually are in a homeostasis state, where a dynamic equilibrium between the different parts guarantees to the being the best allowable performances. Usually, when one of the components is defective or is completely not working, the others adapt to compensate the system. Now, looking at the postural control system, we have that the visual component it is just one of the blocks composing it. There are other proprioceptive and exteroceptive paths that affect the postural sway. It is sensible to argue that when the visual condition is missing (i.e. the closed-eyes condition), the relevance of the other components that regulate the process of equilibrium will be strengthened. Therefore, a study of both the visual conditions together, could take advantage of the different information. For this reason a quaternion-based model of postural sway was investigated.

Here arises the last research question we posed.

Is the postural control system described in a more complete way relying on the information of both open-eyes and closed-eyes time series, therefore with a quaternion representation, or a complex representation is able to obtain the same information?

To answer the question, we decided to confront the results from a complex- and quaternion-based representation of the data, obtained through a prediction scheme recurring to adaptive filtering (Figure 10).

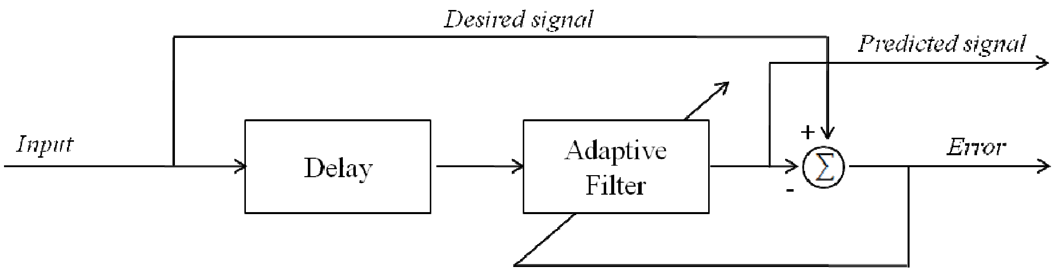


Figure 10 - Scheme for the prediction of a signal using an adaptive filter

The proposed classical scheme for prediction was implemented using four types of adaptive filters:

- Complex Least Mean Square (CLMS)
- Augmented Complex Least Mean Square (ACLMS)
- Quaternion Least Mean Square (QLMS)
- Augmented Quaternion Least Mean Square (AQLMS)

Those filters, defined in (15), (16) and (17) belong to the LMS family, part of the stochastic gradient descent methods. The augmented versions of CLMS and QLMS take into account the so-called augmented statistics valid for complex and hypercomplex random variables. As extensively described in paragraph 9.2, the aim is to consider the pseudo-covariance of the signal, avoiding to assume its circularity. To assess whether the CoP is a circular o non-circular signal, we recurred to the statistical test for circularity of non-gaussian random variables, presented in 8.8.2.

Before the elaboration, each time series was normalized to zero mean to improve the training and test processes. The variances were not normalized to avoid unwanted effects on the circularity/noncircularity of the signal.

To evaluate the performances of the different adaptive filters, we recurred to the prediction gain used in (17):

$$R_p = 10 \log_{10} \frac{\sigma_x^2}{\sigma_e^2} \qquad [\mathrm{dB}]$$

where σ_x^2 is the variance of the input and σ_e^2 the variance of the error (output of the filter minus desired response). We calculated the gain at the steady state. In the case of the quaternion filter were calculated the prediction gains for the OE and CE components separately, to allow a direct comparison with the complex counterpart.

To avoid the possible effect of the cognitive component on the single trial, we decided to analyse the mean time series of a subject of the first dataset. Therefore, we calculated the mean value over all the ten trials, obtaining four mean time series: AP and ML, in both OE and CE conditions. Doing so, we implicitly assumed the ergodicity of the signal. Even if to be proven, this was the only sensible approach to join OE and CE time series. Moreover, in (3) it is explained that Collins and De Luca for the Stabilogram Diffusion Analysis made the same assumption of ergodicity to calculate the mean displacement of the time series.

References in the chapter Algorithms:

Adaptive filters 8.1
Circularity test 8.2

5. Results obtained

1. *Evaluation of the complex nature of CoP*

Anteroposterior (AP) and mediolateral (ML) time series of the CoP were studied together creating a complex signal with AP as the real part, and ML as the imaginary part.

The null hypothesis was that the signal has a bi-variate nature. The alternative hypothesis was that the signal has a complex-valued nature. The significance level of the Kolmogorov-Smirnov test was set to 5%.

Below are presented two tables with the numbers of rejection of the null hypothesis, in the case of time series of 3 and 15 seconds (it was preferred to study shorter time series to reduce the computational time), and under the OE and CE conditions.

The first one presents the total amount of rejected null hypothesis

Condition	3 seconds	15 seconds
CE	94/110	102/110
OE	92/110	100/110

Table 1 - Total results of the complex nature test

The second one presents the amount of rejected null hypothesis per subject /10

Subject number	CE – 3 s	CE – 15 s	OE – 3 s	OE – 15 s
1	9/10	10/10	7/10	10/10
2	8/10	10/10	8/10	9/10
3	10/10	9/10	10/10	10/10
4	7/10	9/10	8/10	8/10
5	9/10	9/10	10/10	10/10
6	9/10	10/10	8/10	8/10
7	8/10	8/10	9/10	9/10
8	9/10	10/10	9/10	9/10
9	7/10	9/10	7/10	10/10
10	10/10	9/10	8/10	8/10
11	8/10	9/10	8/10	9/10

Table 2 - Results per subject of the complex nature test

From the two tables can be argued that the CE and the OE time series share a complex nature in the same way (i.e. similar number of rejected cases), that longer time series are considered complex more easily, and that for long time series for almost all the subjects is preferable a complex representation of the data.

After all, the results answered our research question, demonstrating that anteroposterior and mediolateral time series of the CoP share information needed to describe in a more complete way the postural control system.

2. *Time lag with Average Mutual Information*

In Figure 11 are represented the trends of the obtained Average Mutual Information (AMI) functions.

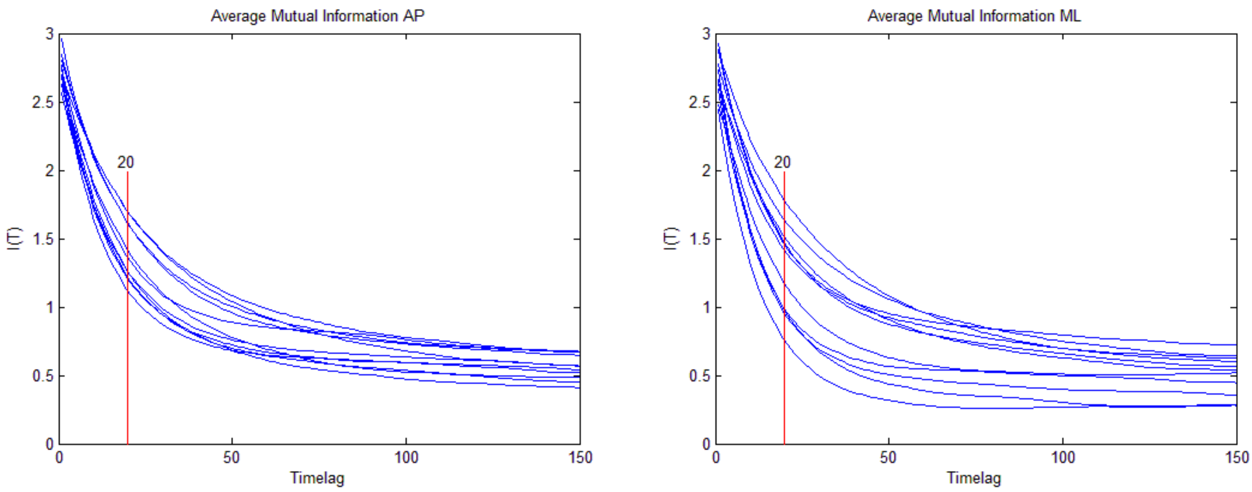


Figure 11 - AMI function for the AP time series (left) and ML time series (right)

Each graph represents the mean of the AMI function for each of the subjects. The mean value was computed by processing together the time series recorded in the closed eyes condition, with the ones recorded in the open eyes condition.

It is established in literature (19), pp. 28-37, to choose the time lag corresponding to the first minimum of the Average Mutual Information. That value represents a delay for which the correlation of the signal and the shifted copy is low enough to reduce redundancy of information, but high enough to avoid incorrelation of the two time series.

Two time delays were selected, for both AP and ML time series, because of their analogous trend. The first one, used in (22) with the same dataset, was $\tau = 0.05\ s$ (it will be used as a reference). The second was $\tau = 0.20\ s$.

The latter choice was due to a qualitative analysis of the COP frequential characteristics. Indeed, it is well known that postural control is characterized by frequencies mainly minor than ~5 Hz (5). Choosing $\tau = 0.20\ s$, means to create delay vectors that correspond to windowed part of the original time series sampled at 5 Hz. Thus, it is reasonable to choose that time delay. A shorter time delay would suffer of the oversampling of the time series.

3. *Embedding parameters with DVV scatter diagrams*

To reduce the computational load, only 30 seconds after the first 10 seconds of one of the COP time series was studied. That choice allowed us to compute the DVV using all the delay vectors (DVs). The other parameters used for the computation were $n\_alpha = 25$ and $n\_sig = 3$ (for a reference on those parameters, see paragraph 8.2).

The results for three embedding dimension and the two time lags chosen with the AMI are reported in Figure 12.

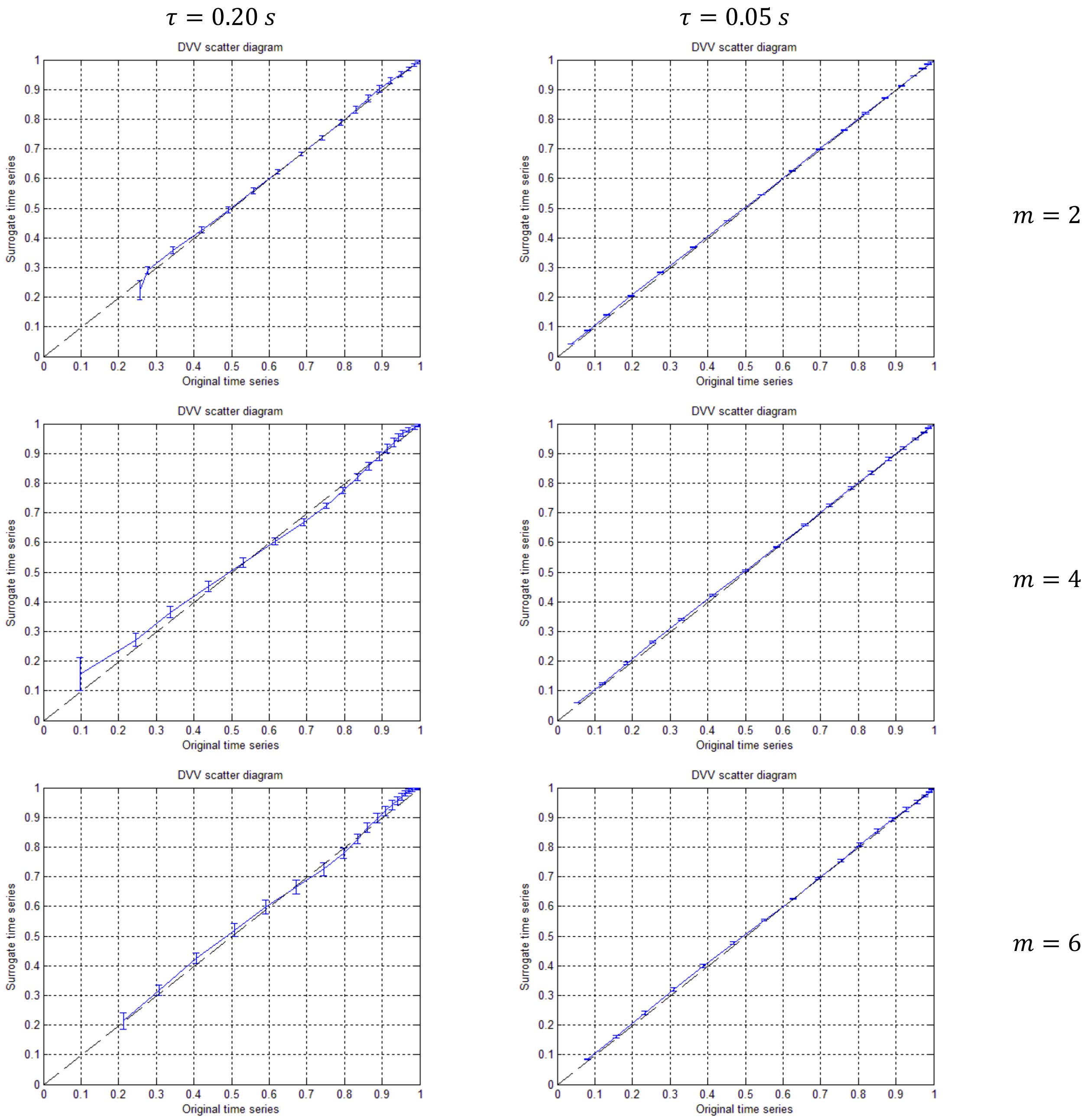


Figure 12 - Dvv scatter plots at different time delays and embedding dimensions

As shown by the DVV scatter diagrams, that look really similar, it is confirmed that DVV is a methodology robust to the change of parameters (25).

However the results obtained for those parameters were misleading, because of a noticeable low degree of non linearity in the COP time series. Because different embedding dimensions gave similar results, two new time delays were chosen from the AMI plots. In the next table are reported the relative DVV scatter diagrams.

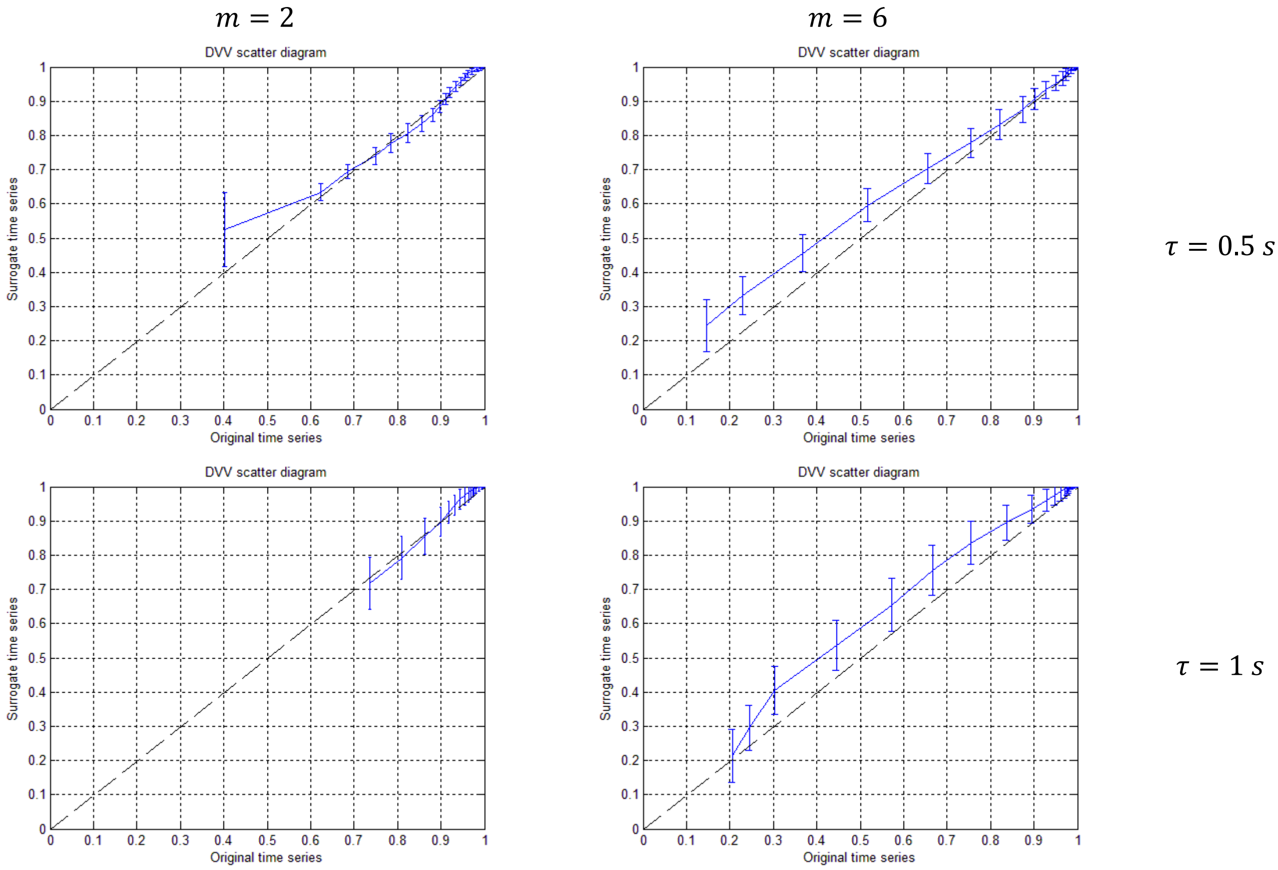


Figure 13 - Dvv scatter plots with higher time delays

An increase in τ or an increase in m corresponds to an increase in the non linearity degree. To certify this trend, a quantitative study was made computing the mean values for each subject's time series of two DVV scatter diagram parameters presented in chapter 4:

- Maximum deviation from linearity (index of nonlinearity)
- Minimum normalized variance (index of stochasticity)

OE and CE time series were studied together.

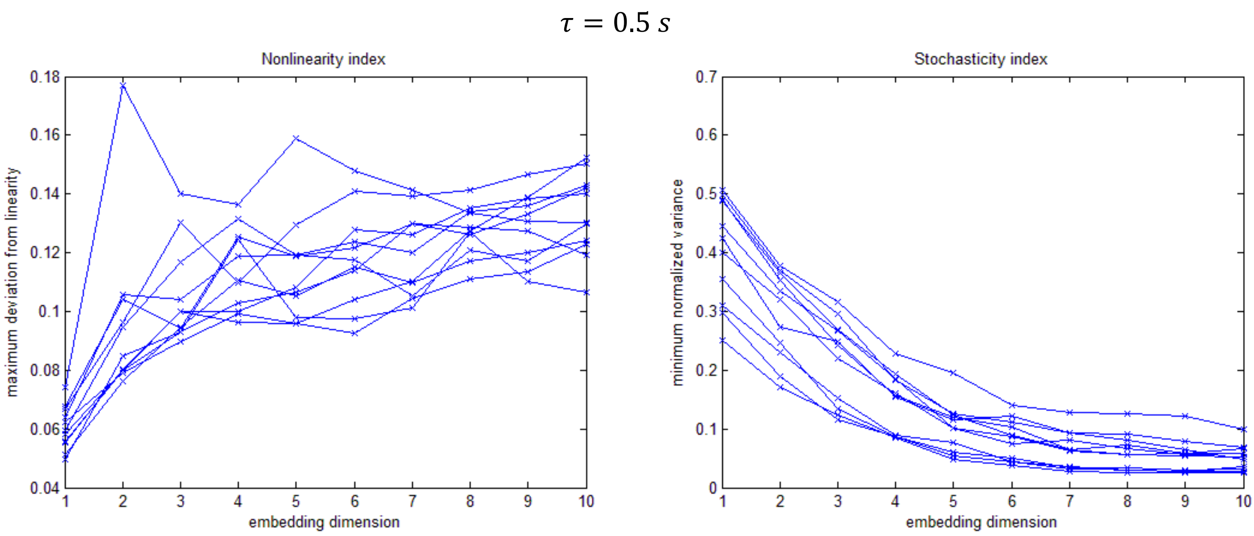


Figure 14 - Nonlinearity and stochasticity indexes as a function of the embedding dimension

The trends confirmed the augmentation of the nonlinearity degree and the determinism together with the embedding dimension. The latter effect is well known, and it is usually due to the displacement of the noisy samples in points far from the attractor of the dynamics. The former effect is the one used to select the proper embedding dimension.

The DVV is a methodology created to assess the degree of determinism and nonlinearity of a time series. We already know of the nonlinear nature of the COP (35).Therefore, we used this methodology as a reference to find a proper embedding dimension. Because of the slow rising trend, $m = 4$ was chosen as the right compromise between the nonlinearity explanation and the computational load. Together with that, it is accepted not to exceed the embedding dimension because of the possible introduction of false determinism (as noticeable from the decreasing trend in the stochasticity index). That value fits the ones used normally in the postural sway literature (22), (7), (33).

A high inter-subject variability was noticed.

To confirm the choice of the time delays, the two parameters were studied as a function of τ.

$$m = 4$$

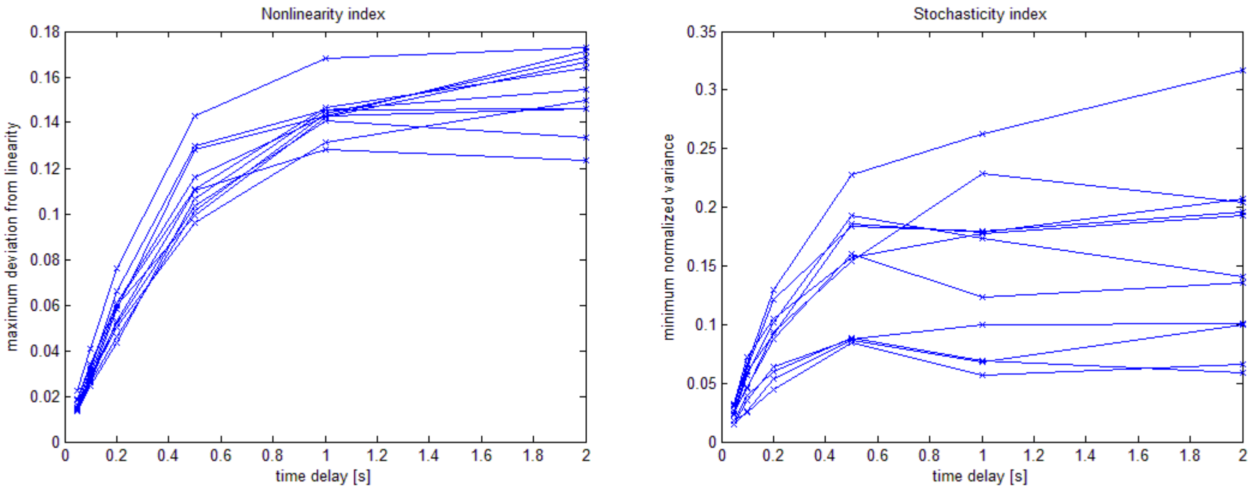


Figure 15 - Nonlinearity and stochasticity indexes as a function of the time delay

In the plots it is shown a trend similar to the one found in the AMI curves. After $\tau = 0.5\ s$, there is a plateau for stochasticity, followed for $\tau = 1\ s$ by nonlinearity.

Those results confirmed the necessity not to follow the standard time delays used in literature, usually near to $\tau = 0.05\ s$, but to explore longer delays, as the AMI function suggested (the real minimum of the curves was near to $\tau = 0.5\ s$). The final set of parameters that was used for all the following studies was:

$$m = 4, \tau = [0.2, 0.5]\ s$$

4. Radius through Recurrence Analysis

In the following, the results of $\%Recurrence$, $\%Determinism$ and $\frac{\%Determinism}{\%Recurrence}$ are shown for $\tau = 0.2\ s$ and $\tau = 0.5\ s$. In both the cases, an embedding dimension of four and a minimum line length of two were used. The minimum line length it's a parameter used for the computation of $\%Determinism$, see paragraph 8.0 for references. The results were obtained averaging the RQA parameters of the AP detrended time series for each subject, with the conditions OE and CE analyzed separately. Only 30 seconds of the signal were studied (neglecting the first 10 seconds recorded), because of the memory usage that is needed for the creation of a recurrence plot (huge matrixes are created, and a high number of samples can easily let the elaborating system run out of memory). The radius r (i.e. the threshold) was multiplied by the standard deviation of the signal before being applied for normalization purposes.

Starting with the closed eyes condition.

$$m = 4,\ \tau = 0.2\ s$$

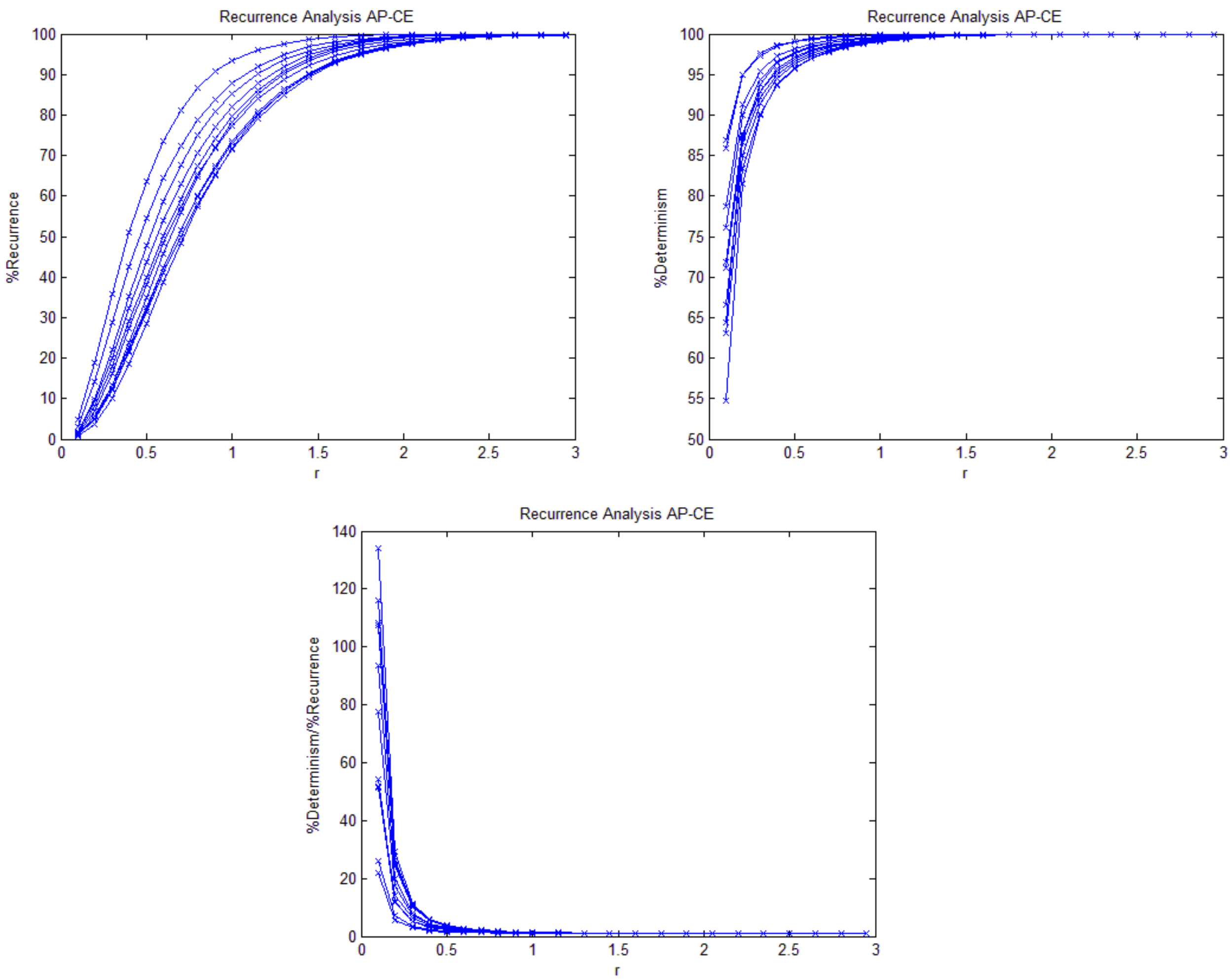


Figure 16 - RQA as a function of the threshold in the CE condition, time delay of 0.2 seconds

As expectable, the determinism degree and the nonlinear autocorrelation increase with the radius. However, they follow different trends. In fact, the degree of determinism rises much quicker than the autocorrelation. When $\%Determinism$ reaches the 100%, all the recurrence points are part of a diagonal. Under that condition, when $\%Recurrence$ rises, all the added recurrence points are part of a diagonal.

$\frac{\%Determinism}{\%Recurrence}$ converges to one for radius bigger than one standard deviation of the signal: the trend is qualitatively similar to the one of $\%Determinism$. That result confirms that the information it takes is redundant (12) when it is not used to reveal changes in the studied system, for instance using windowed analysis.

Now the open eyes condition.

$$m = 4,\ \tau = 0.2\ s$$

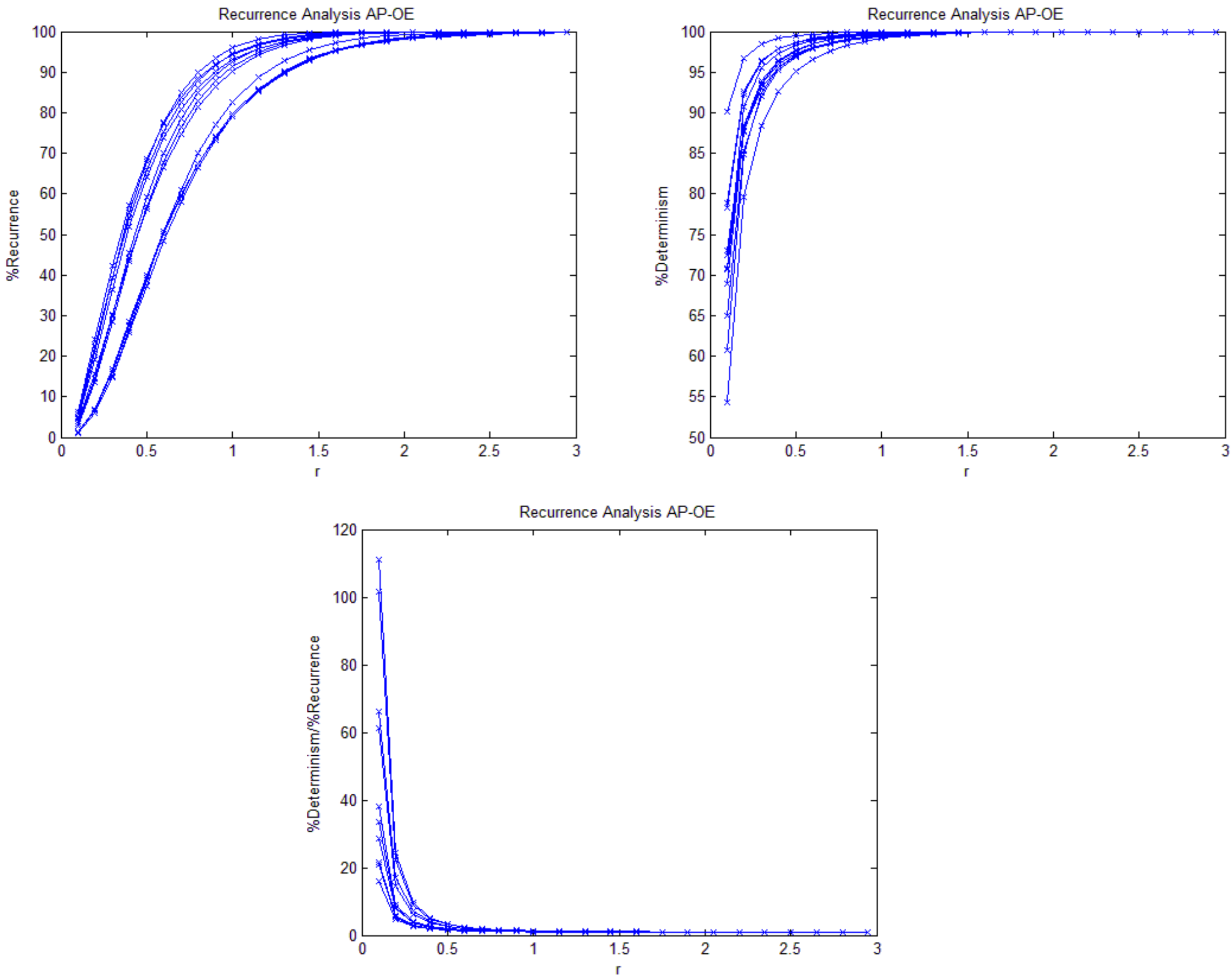
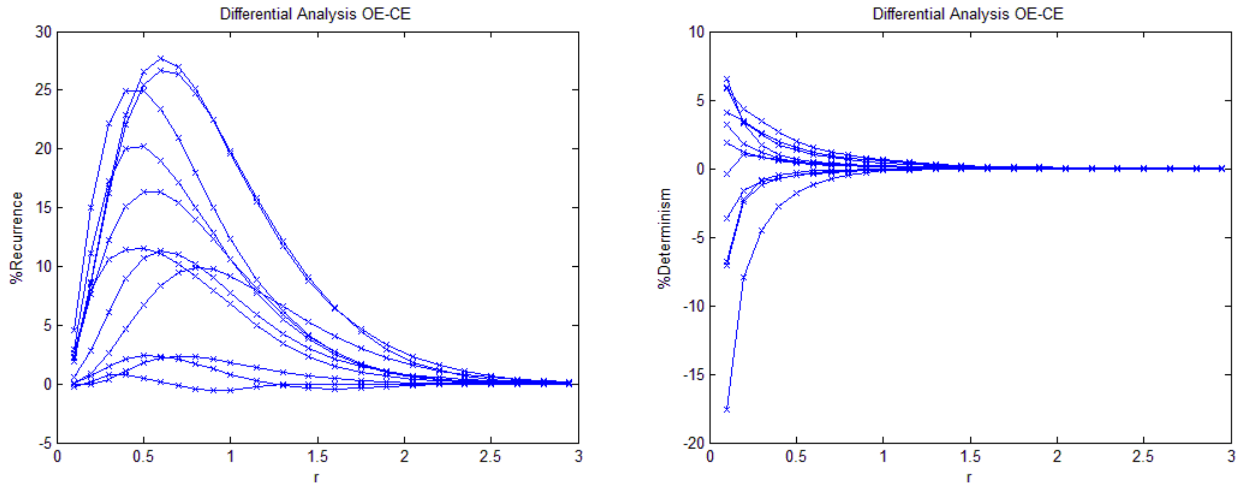


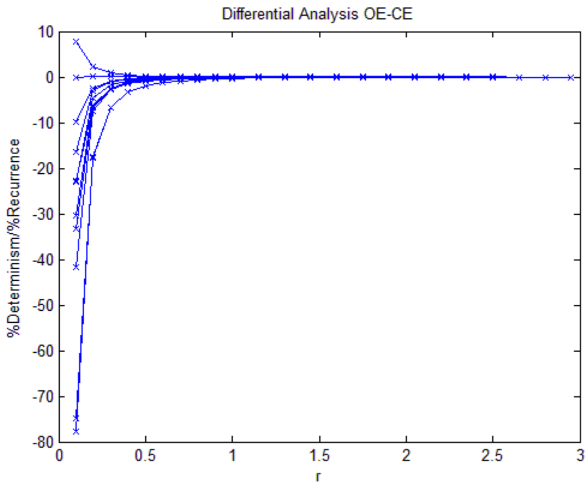
Figure 17 - RQA as a function of the threshold in the OE condition, time delay of 0.2 seconds

The obtained result seems comparable for the shape of the curves and the values with the one in the CE condition.

To better analyze the differences, a differential analysis between the two cases was made.

$$m = 4,\ \tau = 0.2\ s$$





Subject	$\%Rec$	$\%Det$	$\frac{\%Det}{\%Rec}$
1	0.748	0.380	0.502
2	9.146	0.376	1.812
3	0.848	0.473	0.901
4	4.620	0.636	0.631
5	11.227	0.819	**4.657**
6	4.368	**1.584**	2.357
7	4.451	0.258	1.320
8	0.229	0.540	0.021
9	6.604	0.823	1.508
10	**11.472**	1.072	4.528
11	7.527	0.223	1.939

Figure 18 - RQA differential analysis as a function of the threshold, time delay of 0.2 seconds. In the table (bottom-right) are reported the mean values of the curves. The number in bold represent the higher differences

All the three graphs show that the inter-subject trends are different. There are subjects like number 10 that express a high difference between the two conditions in $\%Recurrence$. Others like number 8 show almost equal trend in both the conditions. It is important to notice that the subjects with highest difference in $\%Recurrence$ are not the ones with the highest difference of $\%Determinism$ or $\frac{\%Determinism}{\%Recurrence}$.

In general, OE is more nonlinearly autocorrelated than CE, and is more deterministic for small radius (it is not just for three subjects).

The same analysis was made for the case $\tau = 0.5\ s$. Starting with the CE condition.

$$m = 4,\ \tau = 0.5\ s$$

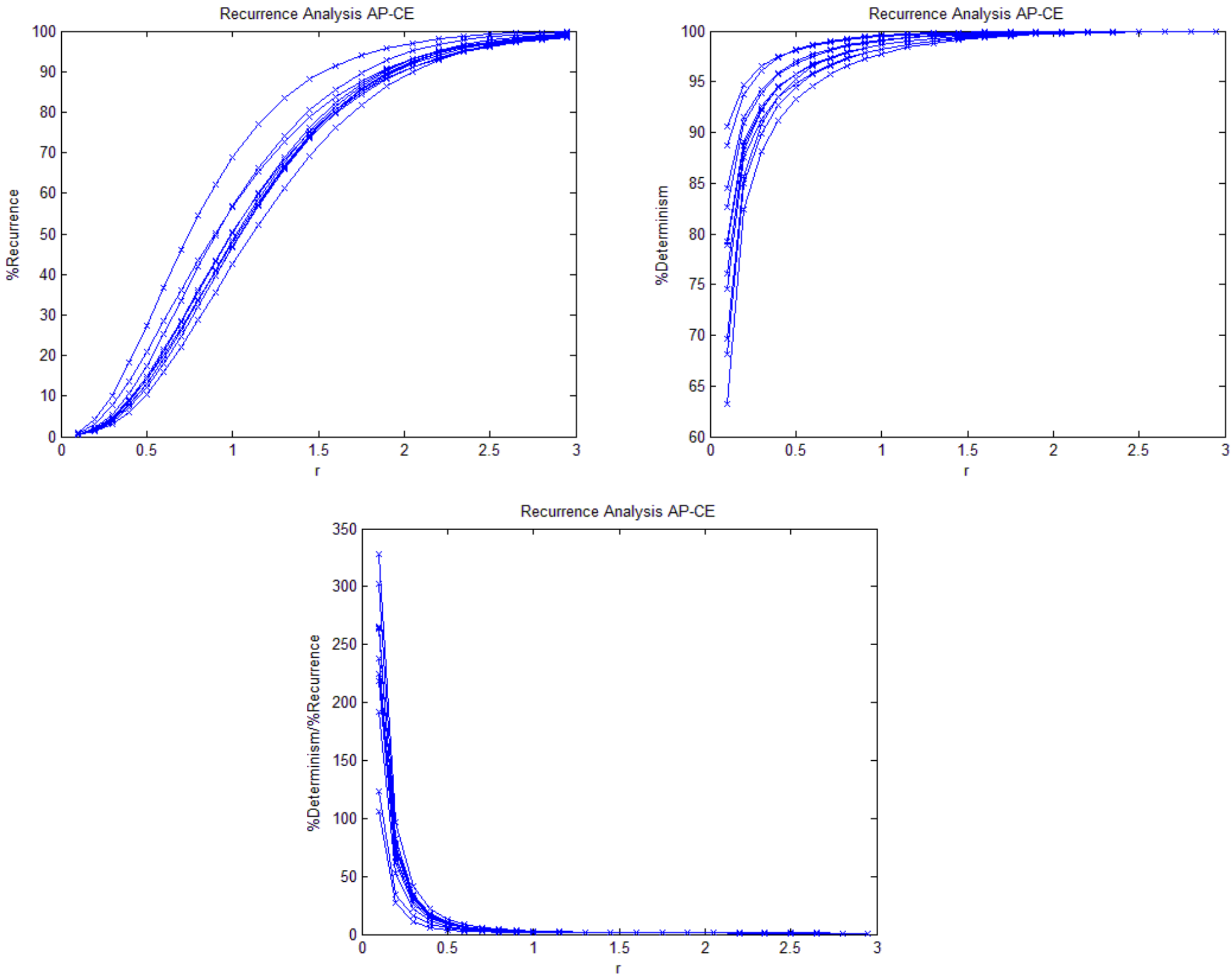


Figure 19 - RQA as a function of the threshold in the CE condition, time delay 0.5 seconds

Now the rising trend for both $\%Recurrence$ and $\%Determinism$ is slower, compared with the case $\tau = 0.2\ s$. Only $\frac{\%Determinism}{\%Recurrence}$ seems to follow a trend similar to the one obtained in the case $\tau = 0.2\ s$.

Then the open eyes condition was elaborated.

$$m = 4,\ \tau = 0.5\ s$$

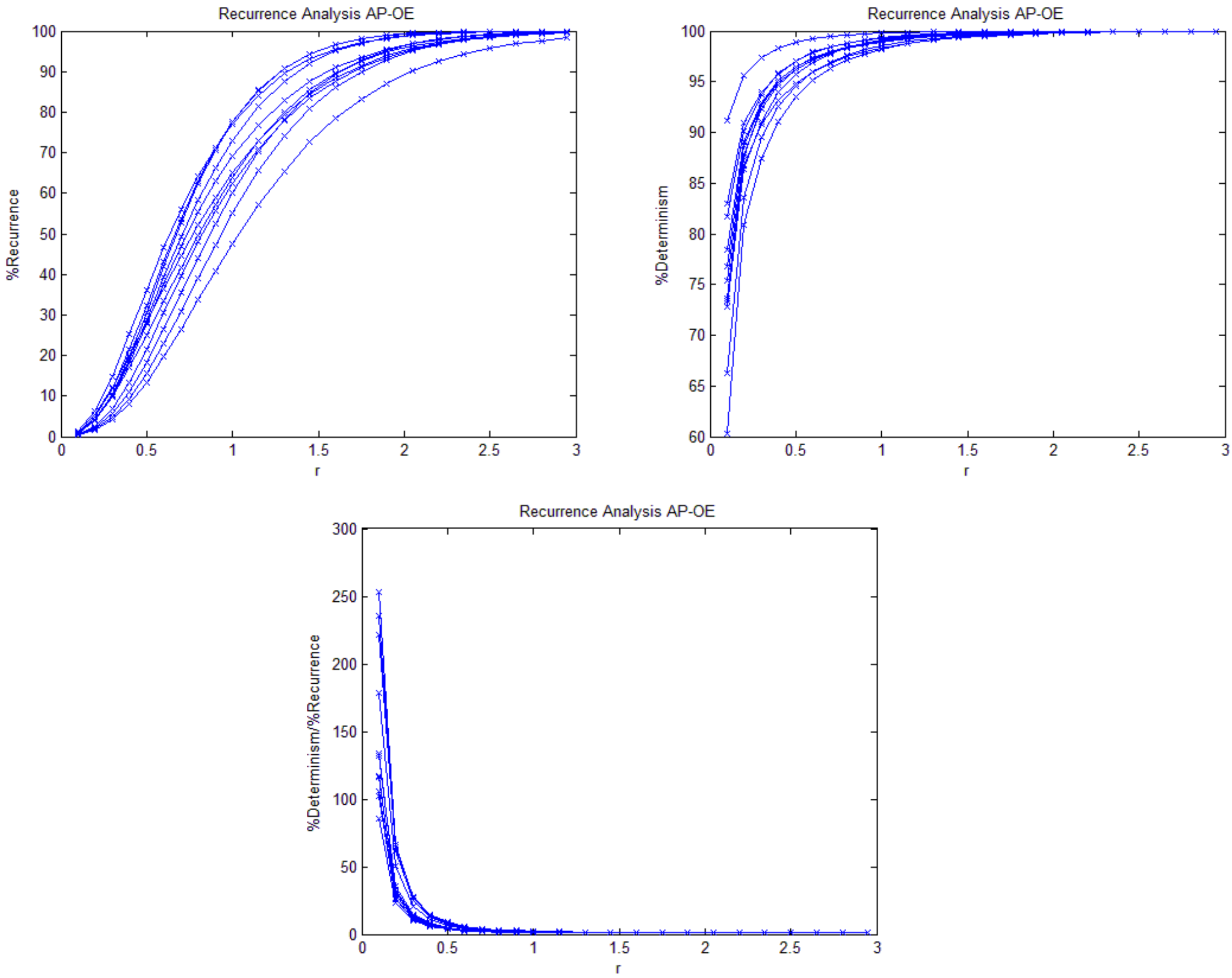
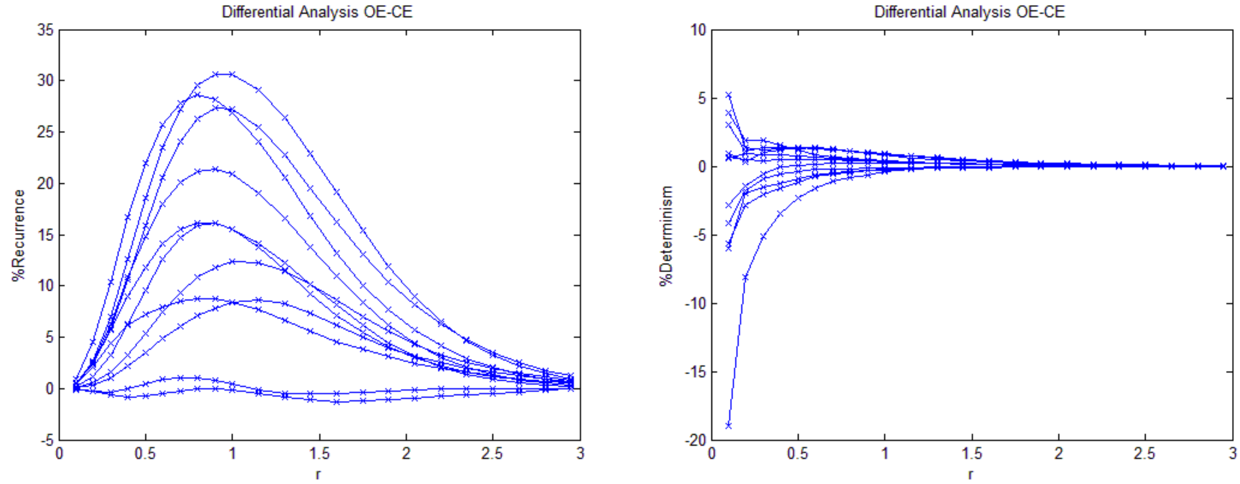
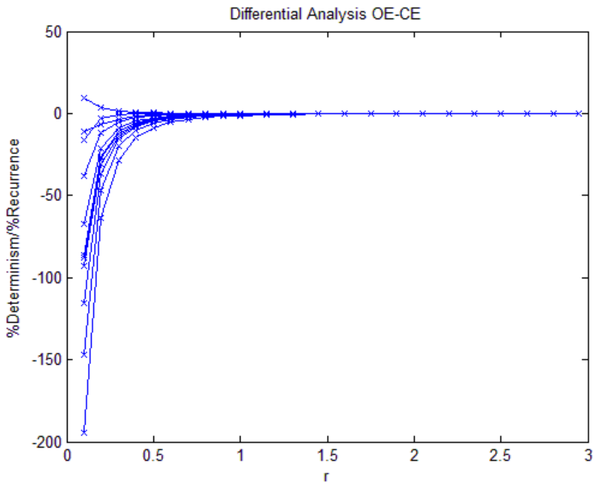


Figure 20 - RQA as a function of the threshold in the OE condition, time delay 0.5 seconds

Again, it was needed to make a differential analysis to understand the differences of the visual conditions.

$m = 4,\ \tau = 0.5\ s$





Subject	$\%Rec$	$\%Det$	$\frac{\%Det}{\%Rec}$
1	4.225	0.380	1.132
2	13.697	0.307	7.968
3	0.544	0.628	0.870
4	4.677	0.620	2.594
5	**15.258**	0.609	**13.697**
6	6.048	**1.816**	6.074
7	7.174	0.245	4.616
8	0.353	0.695	0.725
9	7.753	0.705	6.561
10	13.468	0.846	10.158
11	10.200	0.297	5.870

Figure 21 - RQA differential analysis as a function of the threshold, time delay of 0.5 seconds. In the table (bottom-right) are reported the mean values of the curves. The number in bold represent the higher differences

As seen before, OE is more nonlinearly autocorrelated than CE. Now five subjects show that their CE time series are more deterministic than the OE.

To make an easier evaluation of the cases $\tau = 0.2\ s$ and $\tau = 0.5\ s$, the same tables shown before were represented and compared.

$\tau = 20$				$\tau = 50$			
Subject	$\%Rec$	$\%Det$	$\frac{\%Det}{\%Rec}$	Subject	$\%Rec$	$\%Det$	$\frac{\%Det}{\%Rec}$
1	0.748	0.380	0.502	1	4.225	0.380	1.132
2	9.146	0.376	1.812	2	13.697	0.307	7.968
3	0.848	0.473	0.901	3	0.544	0.628	0.870
4	4.620	0.636	0.631	4	4.677	0.620	2.594
5	11.227	0.819	**4.657**	5	**15.258**	0.609	**13.697**
6	4.368	**1.584**	2.357	6	6.048	**1.816**	6.074
7	4.451	0.258	1.320	7	7.174	0.245	4.616
8	0.229	0.540	0.021	8	0.353	0.695	0.725
9	6.604	0.823	1.508	9	7.753	0.705	6.561
10	**11.472**	1.072	4.528	10	13.468	0.846	10.158
11	7.527	0.223	1.939	11	10.200	0.297	5.870

Table 3 - RQA mean values of the differential analysis OE-CE, computed with a time delay of 0.2 seconds (left) and 0.5 seconds (right)

Because of the shape of the curve, it can be argued that the lowest mean scores of $\%Determinism$ own to the subjects that present a higher $\%Det$ in the CE condition (the mean values were taken down by the starting negative values obtained for small radius).

In the following table the ranking of the different subjects was represented.

Rank $\%Rec$	$\tau = 20$	10	5	2	11	9	4	7	6	3	1	8
	$\tau = 50$	5	2	10	11	9	7	6	4	1	3	8
Rank $\%Det$	$\tau = 20$	6	10	9	5	4	8	3	1	2	7	11
	$\tau = 50$	6	10	9	8	3	4	5	1	2	11	7
Rank $\frac{\%Det}{\%Rec}$	$\tau = 20$	5	10	6	11	2	9	7	3	4	1	8
	$\tau = 50$	5	10	2	9	6	11	7	4	1	3	8

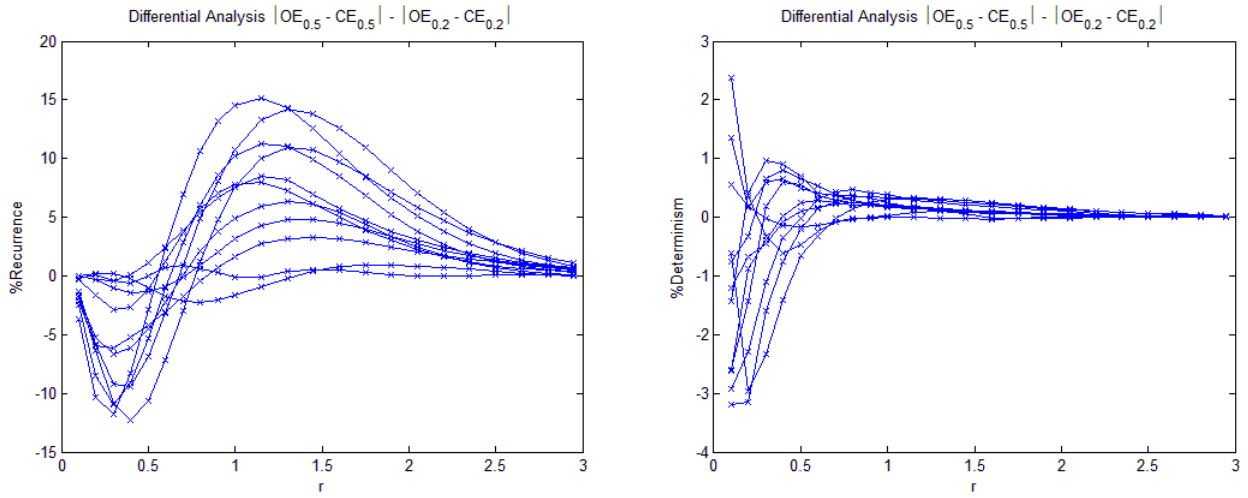
Table 4 - Ranking of the subjects RQA. The higher values are on the left, the lower on the right

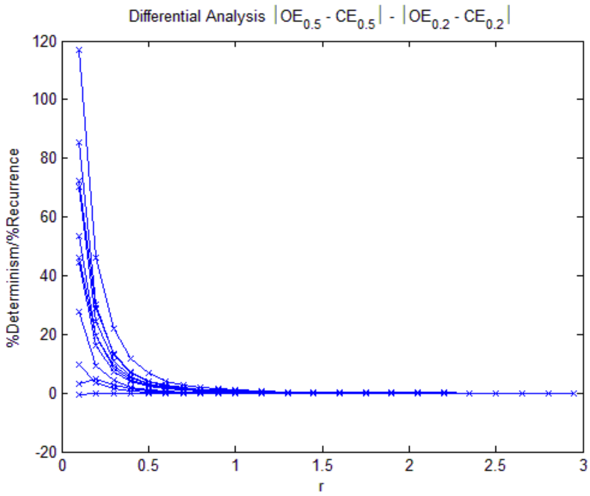
The subjects placed at the same position for both the values of time delays are underlined in yellow. From the table it can be noticed that the ranking is preserved in most of the cases, if a translation of one place is accepted (blue underlining). The most variable results $\frac{\%Determinism}{\%Recurrence}$.

From the graphs shown until now, it can be argued that for small radius the degree of determinism mainly depends on the characteristics of the subject. For higher radius, the determinism in CE and OE conditions gets similar, even if the nonlinear autocorrelation is higher for the OE time series. These results are true for both the time delays. What changes are the score and the order between the subjects.

A differential analysis between the two time delays differential series was made to understand which time delay explains the higher difference between OE and CE conditions.

$\tau = 0.5\ s$ and $\tau = 0.2\ s$





Subject	$\%Rec$	$\%Det$	$\frac{\%Det}{\%Rec}$
1	3.547	0.060	0,629
2	**7.635**	0.253	6,155
3	0.842	0.245	0,045
4	2.464	0.318	1,962
5	7.028	**0.530**	**9,039**
6	2.117	0.235	3,717
7	3.441	0.187	3,296
8	0.254	0.273	0,703
9	3.359	0.439	5,052
10	6.332	0.468	5,630
11	5.796	0.220	3,930

Figure 22 - Differential analysis of the absolute differential series OE-CE for the evaluation of the time delay differences

$\%Rec$urrence shows that time series for $\tau = 0.2\ s$ holds more information about the differences between OE and CE, in an interval $r = [0.01,\ 0.7]$. For the other values of r, it is $\tau = 0.5\ s$ that shows a higher difference of nonlinear autocorrelation between the open eyes and closed eyes conditions. The degree of determinism is more variable, even if the dynamics is mainly compressed in the range $r = [0.01,\ 1]$. For higher values of the radius, the difference in the determinism is almost zero.

The following set of parameters will be used for the Fuzzy Entropy analysis:

$$m = 4, \tau = 20, r = [0.5, 1.5]$$
$$m = 4, \tau = 50, r = [0.5, 1.5]$$

Given the fact that to a higher nonlinear autocorrelation corresponds to a lower degree of entropy, what we do expect is that:

- For both the parameter's sets, the OE entropy will be lower than the CE for the subjects with an high differential $\%Recurrence$
- The FE ranking of the subjects will be the same one of $\%Recurrence$, for both the time delays

5. Fuzzy entropy of the whole time series

5.1 Initial evaluation

The FEs score of the subject time series were studied using the values of the threshold identified in the latter paragraph.

Starting from $r = 0.5$, in the following table are showed the mean values and the standard deviations of all the time series for each subject.

Subject	CE-AP $\tau = 0.2\ s, r = 0.5$	OE-AP $\tau = 0.2\ s, r = 0.5$	CE-AP $\tau = 0.5\ s, r = 0.5$	OE-AP $\tau = 0.5\ s, r = 0.5$
1	0.39 ± 0.07	0.34 ± 0.11	0.61 ± 0.09	0.53 ± 0.14
2	0.32 ± 0.03	0.21 ± 0.07	0.53 ± 0.05	0.38 ± 0.12
3	0.35 ± 0.07	0.37 ± 0.07	0.54 ± 0.10	0.55 ± 0.08
4	0.25 ± 0.06	0.24 ± 0.05	0.43 ± 0.10	0.44 ± 0.08
5	0.42 ± 0.10	0.22 ± 0.09	0.63 ± 0.11	0.36 ± 0.13
6	0.39 ± 0.11	0.32 ± 0.09	0.54 ± 0.14	0.49 ± 0.14
7	0.39 ± 0.09	0.34 ± 0.10	0.59 ± 0.13	0.48 ± 0.14
8	0.25 ± 0.05	0.23 ± 0.07	0.45 ± 0.08	0.42 ± 0.11
9	0.37 ± 0.06	0.28 ± 0.07	0.59 ± 0.08	0.49 ± 0.11
10	0.37 ± 0.09	0.21 ± 0.07	0.56 ± 0.11	0.35 ± 0.11
11	0.32 ± 0.04	0.21 ± 0.04	0.51 ± 0.06	0.38 ± 0.07

Table 5 - Values of the Fuzzy Entropy with different visual conditions and time delays, radius 0.5

The differences between OE-CE are preserved for different time delays. That is not true for subject four. The standard deviations, as the mean values are always higher for $\tau = 0.5\ s$ than for $\tau = 0.2\ s$. Usually CE entropy is higher than OE entropy, the only exceptions are subjects with similar values OE-CE, like subjects three and four.

The next table reports the Fuzzy Entropy values for $r = 1.5$.

Subject	CE-AP $\tau = 0.2\ s, r = 1.5$	OE-AP $\tau = 0.2\ s, r = 1.5$	CE-AP $\tau = 0.5\ s, r = 1.5$	OE-AP $\tau = 0.5\ s, r = 1.5$
1	0.09 ± 0.02	0.07 ± 0.03	0.17 ± 0.04	0.12 ± 0.04
2	0.08 ± 0.01	0.04 ± 0.02	0.14 ± 0.02	0.08 ± 0.03
3	0.08 ± 0.02	0.09 ± 0.03	0.13 ± 0.03	0.14 ± 0.03
4	0.06 ± 0.02	0.05 ± 0.01	0.11 ± 0.04	0.13 ± 0.03
5	0.10 ± 0.03	0.04 ± 0.02	0.16 ± 0.04	0.07 ± 0.03
6	0.09 ± 0.03	0.06 ± 0.02	0.13 ± 0.04	0.11 ± 0.04
7	0.09 ± 0.02	0.07 ± 0.03	0.14 ± 0.04	0.10 ± 0.04
8	0.05 ± 0.01	0.05 ± 0.02	0.12 ± 0.02	0.10 ± 0.04
9	0.09 ± 0.02	0.06 ± 0.02	0.15 ± 0.03	0.12 ± 0.04
10	0.08 ± 0.02	0.03 ± 0.01	0.13 ± 0.04	0.07 ± 0.02
11	0.07 ± 0.01	0.04 ± 0.01	0.13 ± 0.03	0.08 ± 0.02

Table 6 - Values of the Fuzzy Entropy with different visual conditions and time delays, radius 1.5

The same consideration of the case $r = 0.5$ can be made. Anyway, the mean values in this case are much lower.

The obtained values confirm that the CE-OE confront gives often different results. Our results are incoherent with the ones proposed by other authors, like(10) and (29), where the SE is higher in the OE case, or (6), where the SE values are equal in the OE and CE conditions. Many are the possible reasons of those differences. Just to give an example, in (10) the differences could be due to the parameters choice ($m = 3$ and $\tau = 0.01\ s$), together with the approach chosen for the elaboration. Indeed, the authors presented the results of the mean values of Sample Entropy calculated over 30 subjects, including AP and ML time series together.

In our opinion, it is important to consider what shown in (13), where a group of athletes and a group of ballet dancers are confronted using RQA. One of the results was:

$$\%Recurrence\ CE\ Athletes > \%Recurrence\ OE\ Athletes$$

$$\%Reccurrence\ OE\ Ballet\ dancers > \%Recurrence\ CE\ Ballet\ dancers$$

Also for other parameters, it was shown that different training leads to different parameters. Another example of a similar study is given in (14). Those results underline the importance to consider the inter-variability of the physical characteristics and the motor skills of the subjects.

5.2 Confront of two subjects

To improve the characterization of the FEs an extensive study of two specific subjects out of the eleven used for the creation of the dataset was made.

Subject 4 and 11 were chosen, because of the most similar and the most different entropy values in the OE-CE conditions, respectively.

The result obtained for them during the estimation of the parameter is represented in the table below.

		Subject 4		Subject 11	
		$\tau = 0.2\ s$	$\tau = 0.5\ s$	$\tau = 0.2\ s$	$\tau = 0.5\ s$
$Nonlinearity$		0.050	0.096	0.051	0.103
$Stochasticity$		0.053	0.086	0.060	0.088
$\%Rec$	OE	83.635	66.086	84.303	71.676
	CE	79.015	61.408	76.775	61.475
$\%Det$	OE	97.876	97.225	98.372	97.927
	CE	97.240	96.605	98.185	98.035

Table 7 - DVV and RQA values for subjects 4 and 11

The nonlinearity and stochasticity indexes were taken from the plots reported in 5.3. Subject 11 expresses a higher nonlinearity and stochasticity. Regarding the RQA parameters, for both the subjects it is true that:

$$\%Rec\ CE\ \tau_{0.5} < \%Rec\ OE\ \tau_{0.5} < \%Rec\ CE\ \tau_{0.2} < \%Rec\ OE\ \tau_{0.2}$$

$$\%Det\ OE\ \tau_{0.2} > \%Det\ CE\ \tau_{0.2} > \%Det\ OE\ \tau_{0.5} > \%Det\ CE\ \tau_{0.5}$$

Subject 11 has an higher $\%Recurrence$ and $\%Detetermism$, compared to subject 4. It is worth to point out that $\%Detetermism$ is coherent with nonlinearity, because when one rises, the other lowers.

In the next graphs the trends of FE for both AP and ML time series were reported, for both $\tau = 0.5\ s$ and $\tau = 0.2\ s$. To uniform the elaboration with the one of the RQA in paragraph 5.0, only 30 seconds after the initial 10 seconds were studied. That allowed us to confront the two results.

$$\tau = 0.5\ s$$

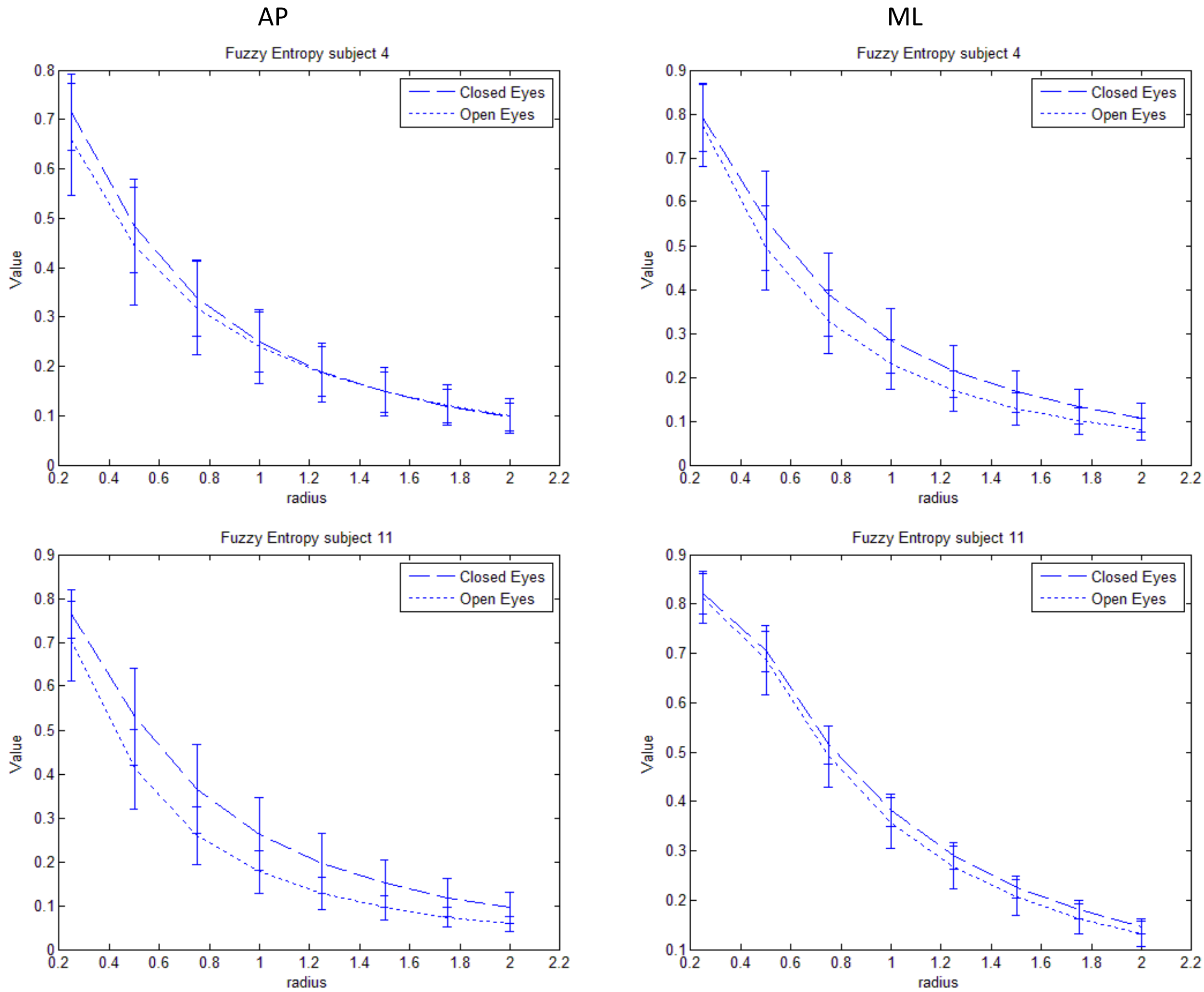


Figure 23 - AP (left) and ML (right) trends of the FE of subject 4 (top) and 11 (bottom) as function of the radius. Both OE and CE conditions were reported. Time delay was 0.5 seconds

It is confirmed that CE entropy is higher than OE entropy. Looking carefully, it can be noticed that subject 11 has mean values higher than subject 4 for radius less than $r = 1$. That is the opposite of what expected, because he had a $\%Recurrence$ higher than subject 11. The reason is explainable thinking that the value of $\%Recurrence$ proposed in Table 7 it is only a mean value of the curves. Indeed, if we look at the curves of $\%Recurrence$, it is possible to notice the coherence of the results (Figure 24).

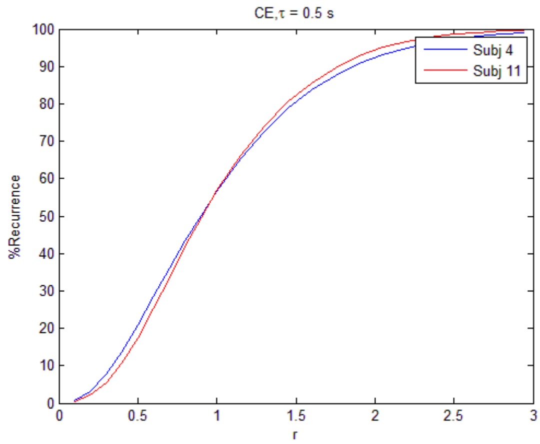


Figure 24 - Detail of the %Recurrence curves of subject 4 and 11

Thus, it is confirmed the inverse proportionality between $\%Recurrence$, index of nonlinear autocorrelation, and the Fuzzy Entropy. Now to the case $\tau = 0.2\ s$.

$$\tau = 0.2\ s$$

AP ML

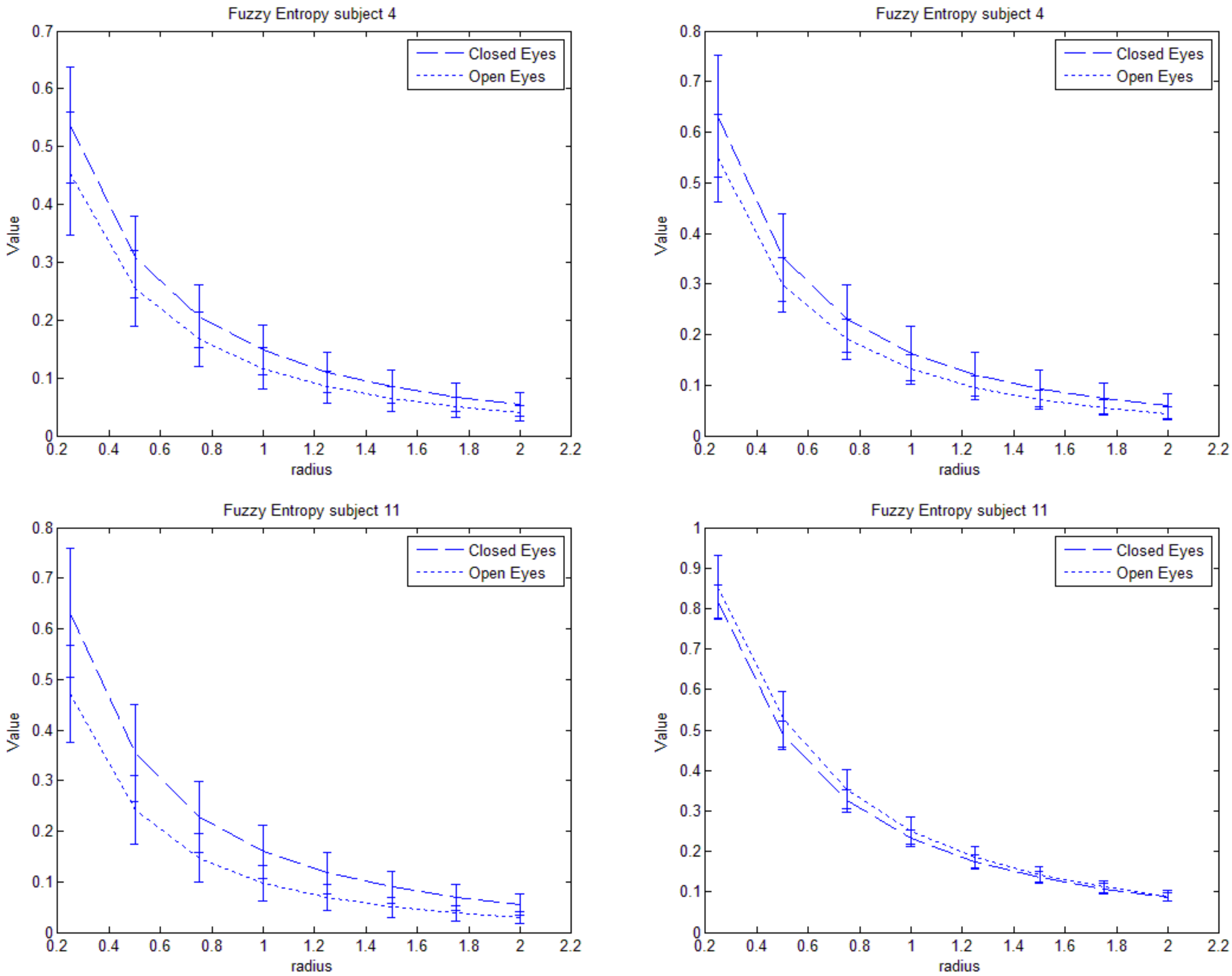


Figure 25 - AP (left) and ML (right) trends of the FE of subject 4 (top) and 11 (bottom) as function of the radius. Both OE and CE conditions were reported. Time delay was 0.2 seconds

The shape of the obtained curves is qualitatively similar to the precedent case. The main exception is the profile of the ML time series of subject 11. For $\tau = 0.2\ s$ OE entropy is higher than CE entropy, while for $\tau = 0.5\ s$ the opposite is true. It is clear that, to assess the relation between FE values and $\%Recurrence$, the same study need to be extended to all the time series.

5.3 Final evaluation

The comparison between FE and $\%Recurrence$ was made using the results obtained in paragraph 5.0, and computing the values of entropy as a function of r. In Figure 26 the results of $\%Recurrence$ are reported again. It was added the relative number of the subject and a different color, just to allow an easier discrimination of the curves.

$m = 4,\ \tau = 0.2\ s$ $\qquad$ $m = 4,\ \tau = 0.5\ s$

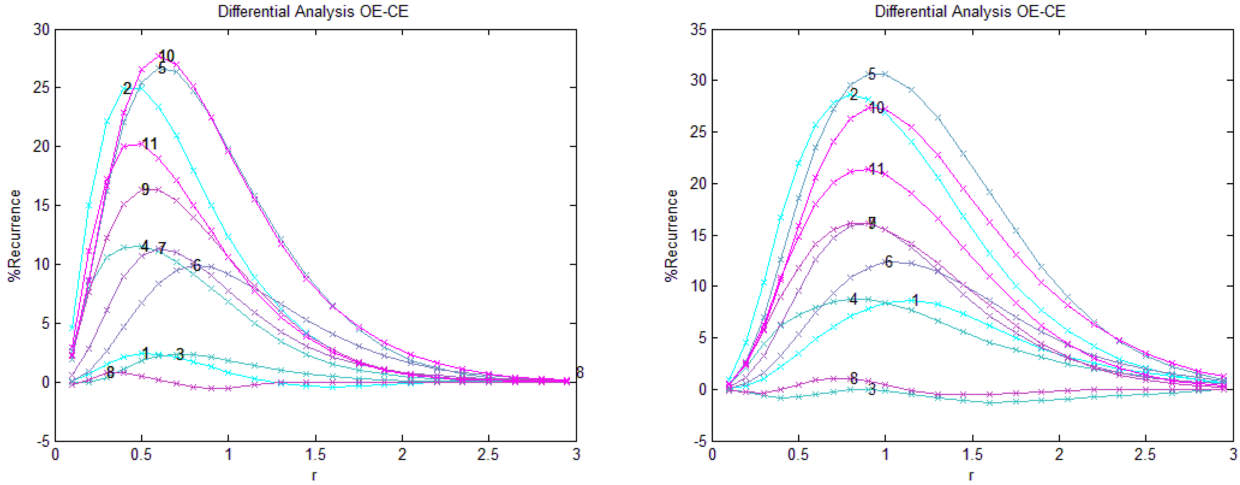


Figure 26 - Figure 18 (left) and Figure 21 (right) trends of %Recurrence. For each profile the number of the relative subject was added

What presented in Figure 26 allows to argue that subject 8 will have FE values higher in the OE case when the threshold is approximately between 0.5 and 1, in the case of $\tau = 0.2\ s$. Together with subject 8, for $\tau = 0.5\ s$ there will be radius values for which subject 3 will have a higher FE in the OE case. Comparing the results with the ones for FE entropy (Figure 27), similar trends can be noticed.

$m = 4,\ \tau = 0.2\ s$ $m = 4,\ \tau = 0.5\ s$

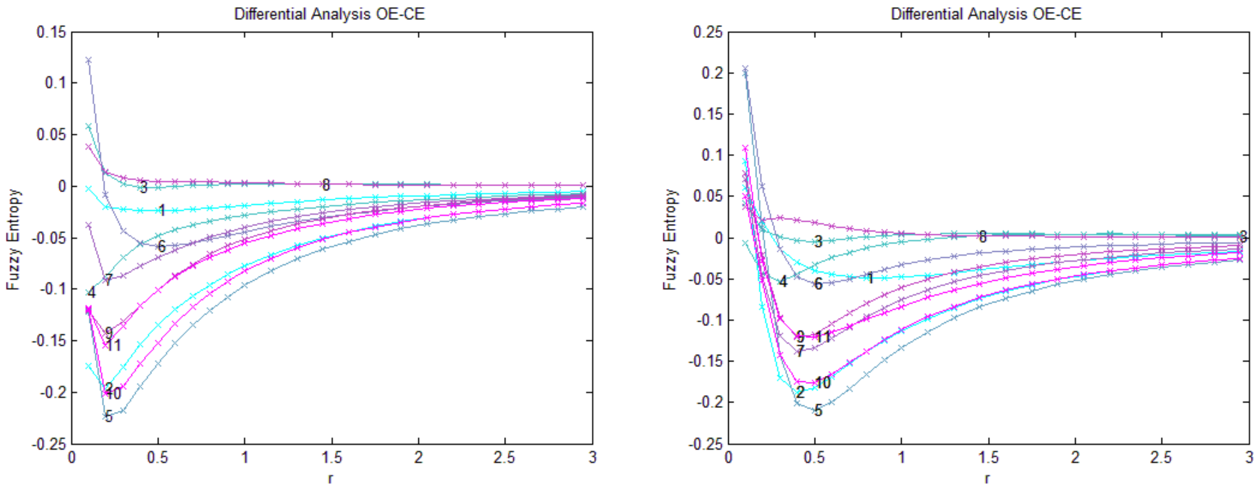


Figure 27 - Trends of the AP Fuzzy Entropy of each subject, as functions of the radius. For each profile the number of the relative subject number was added

The first thing to notice is that for small radii, almost all the subjects present different results than the ones obtained for higher values. In particular, taking subject 6 as example, a radius of 0.2 would lead to a FE higher in the OE case, but a radius of 0.5 would lead to the opposite result. Avoiding small radii, the distinction between the CE and OE conditions for some subjects are radius-dependent, see subject 8 and 3, while for others this is not the case.

For both the time delays, the results achieved through FE are similar to the ones of $\%Recurrence$.

Three are the main features of Figure 26 and Figure 27:

- The peak of the OE-CE difference is function of τ; this is particularly true for $\%Recurrence$
- The order of the subject differences is approximately maintained, for both the methodologies and the time delays
- For radii higher than a certain value, FE and $\%Recurrence$ give coherent results in the discrimination between OE and CE.

For completeness , the same study was extended to the ML time series.

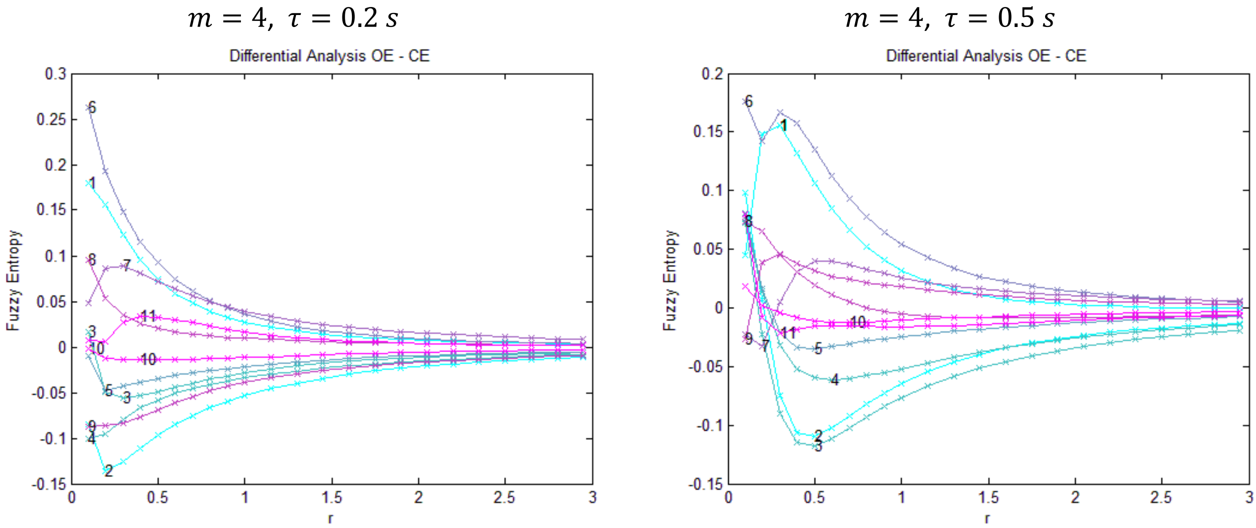


Figure 28 - Trends of the ML Fuzzy Entropy of each subject, as functions of the radius. For each profile the number of the relative subject number was added

The trends in Figure 28 have a similar decay to the one shown for the AP time seris. However, in this case many subjects present a FE higher in the OE case (as seen in Figure 25 for subject 11). Again we find the peaks that translate according to the time delay, and a variable behavior for small values of the radius.

6. *Fuzzy entropy as a function of time*

The purpose of the elaborations was to find temporal intervals where the subjects focused their attention on the postural control. Two kinds of studies were conducted:

- With overlapping windows
- Sequential

Starting with the former case, in are reported the result for $m = 4, r = [0.5, 1.5]$ and $\tau = [0.2, 0.5]\ s$ of subject 11, the one that showed the higher differences in the entropy for those values of the radius. The chosen windows length was 350 samples (i.e. 3.5 seconds) for both the time delays. The overlap was of 349 samples.

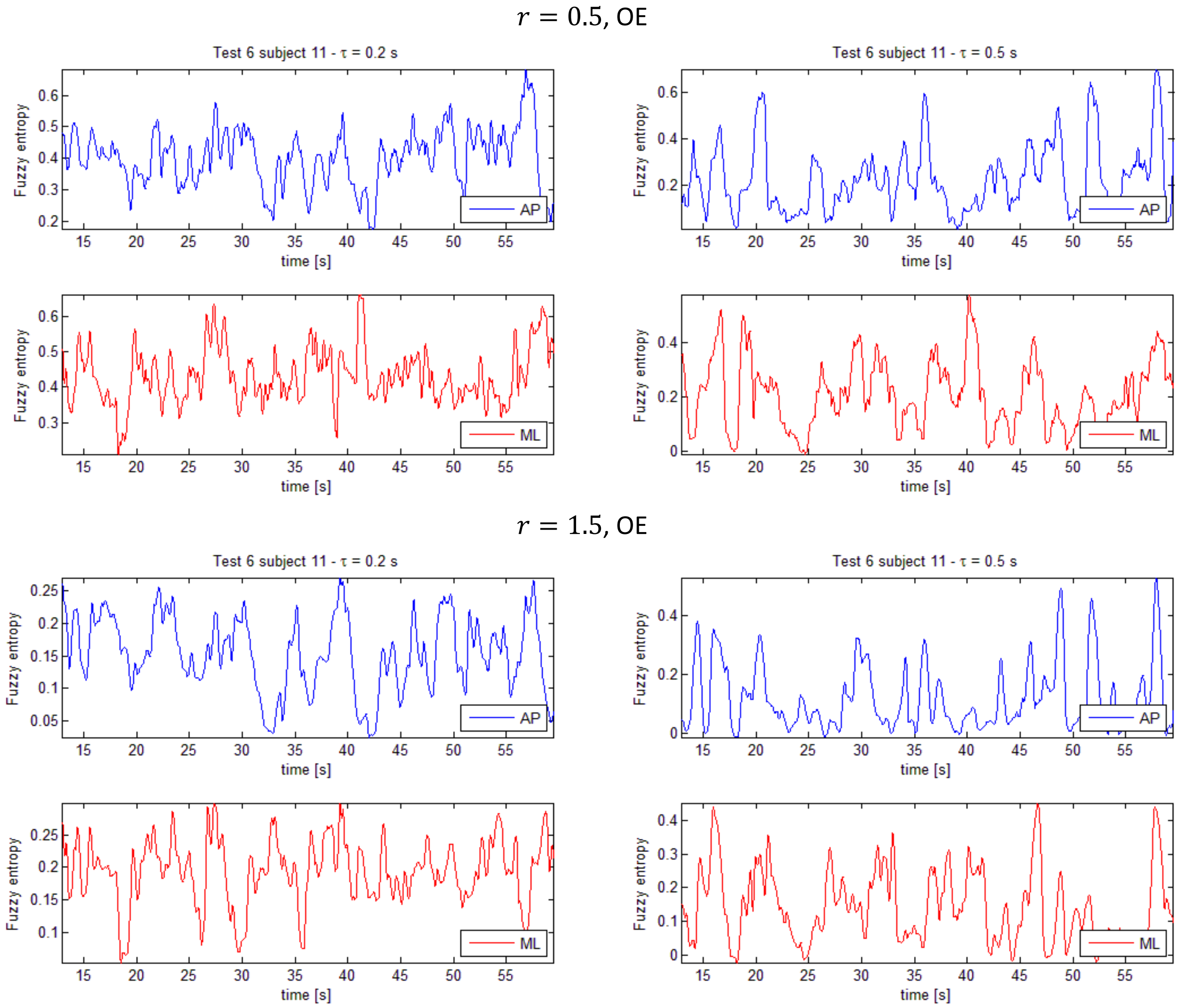


Figure 29 - Subject 11, AP and ML time series of FE obtained with windowed part of the signal (350 samples each)

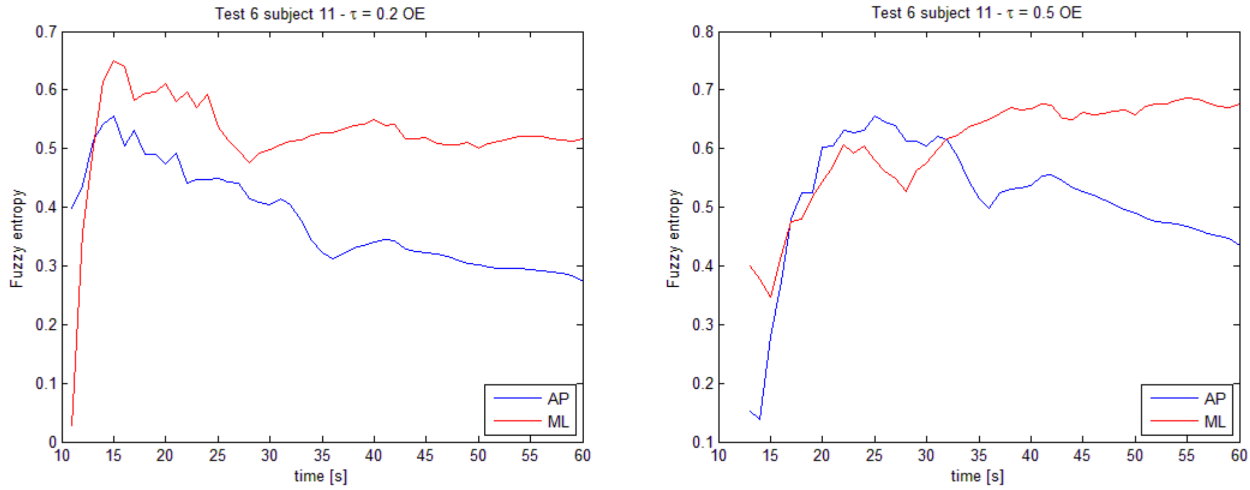
The behavior we expected was a bi-stable one, assuming that the "attention-on" and "attention-off" had two values of FE more or less fixed. Only the case $\tau = 0.5\ s$ recalls that trend, but the number of changes is high, maybe too much and in too less time. It is difficult to argue whether those oscillations are really describing the mind control of the subject, or not. Furthermore, the AP and ML spikes are almost uncorrelated (Table 8), while we argued that the cognitive component should have a coherent effect on both the AP and ML time series (i.e. correlated fluctuations).

	$\tau = 0.2\ s$	$\tau = 0.5\ s$
$r = 0.5$	0.00	0.16
$r = 1.5$	0.06	0.24

Table 8 - Correlation coefficients between AP and ML time series, windowed case

The sequential analyses were reported in Figure 30.

$r = 0.5$, OE



$r = 1.5$, OE

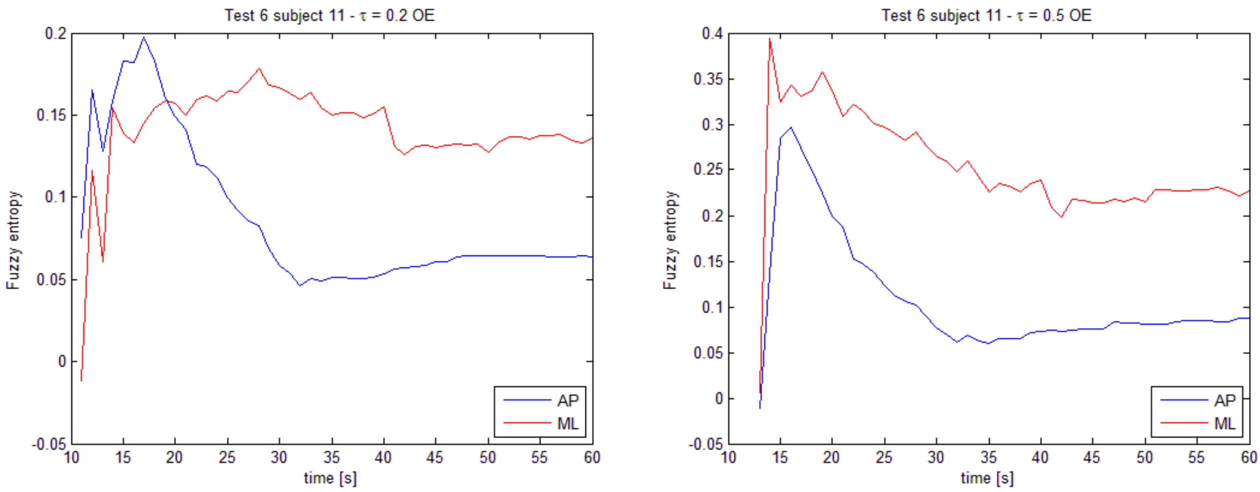


Figure 30 - Subject 11, AP and ML profiles of the sequential FE

Again, both AP and ML profiles are quite different (see Table 9), especially in the initial part. The FE in the ML case is always higher than in the AP case. That features was not noticed for the windowed FE. More or less in all the presented graphs, both AP and ML have a transient and then a sort of plateau. Because of the length of the transient and the property of FE to handle short time series, it must be due to the intrinsic characteristic of the time series (thus, it is not related to a numerical problem).

	$\tau = 0.2\ s$	$\tau = 0.5\ s$
$r = 0.5$	0.24	0.41
$r = 1.5$	0.01	0.77

Table 9 - Correlation coefficients between AP and ML time series, sequential case

The impossibility to understand the cognitive involvement of the subjects led us to the creation of a new dataset, expressly created with that purpose.

7. Study of a specific dataset

7.1 Embedding parameters

To extend the approach used for the first dataset, the AMI function was studied, together with the dependence of the FE to the radius. In Figure 31 the relative graphs are reported.

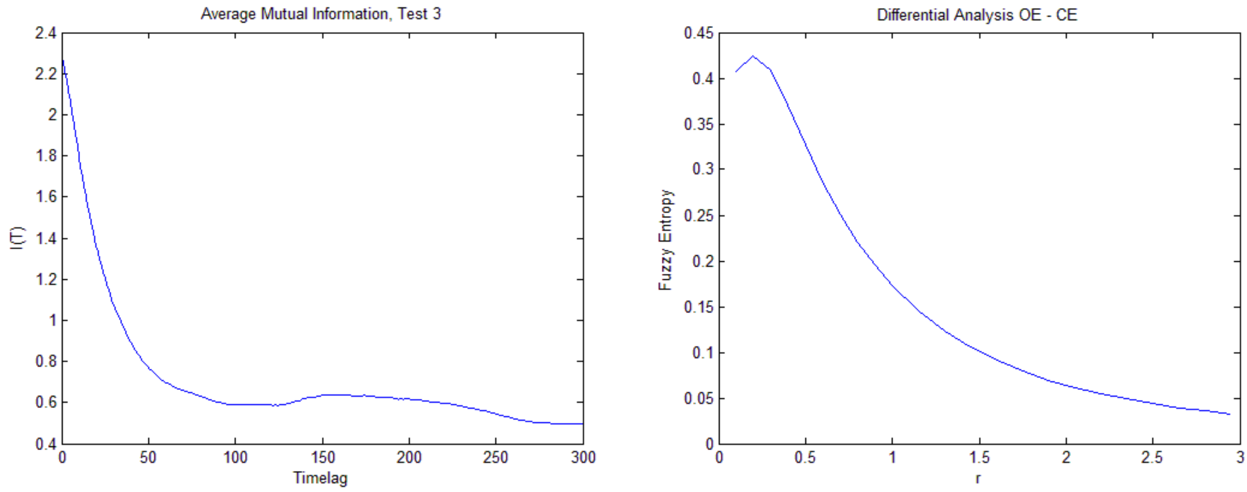


Figure 31- AMI function (left) and FE differential analysis (right) for the choice of the proper parameters

In the case of the AMI function, test 3 was studied because it held the information of both the CE and OE conditions. The FE graph was made using test 1 and 2. Only the AP time series were used.

It can be noticed that the AMI has a trend similar to the one presented in Figure 11. The values are different because of the different sampling frequency (140 Hz for this dataset, 100 Hz for the other one). Therefore, it is reasonable to adopt the same values found previously, considering that $\tau = 0.2\ s$ is placed at 28 samples, and $\tau = 0.5\ s$ is placed at 70 samples.

The embedding dimension was directly extended to this dataset because it was proven during the precedent elaborations that it is not a sensitive parameter.

The study of the proper radius supports the theory of the dependence of the results to the physical features of the subjects. Indeed, in this case the FE is higher in the OE condition.

7.2 Dynamical study of the visual condition

The elaborated parameters were:

- Standard deviation, index of variability
- Sway-path length, index of curviness
- Fuzzy Entropy, index of complexity
- Hurst exponent, index of scaling behavior

The obtained results are directly compared with the ones achieved in the article describing the effect of the attention on the CoP (10). That study was made averaging the results of 30 different subjects, taking together AP and ML time series. Being a inter-subject study, we think that It can be a useful reference to our results. The single-task condition involved just staying in OE or CE condition. The dual-task involved a cognitive exercise to reduce the attention of the subjects on the postural control.

Standard deviation

As reported in Figure 32, the standard deviation of the time series is a sensitive parameter to the eyes condition.

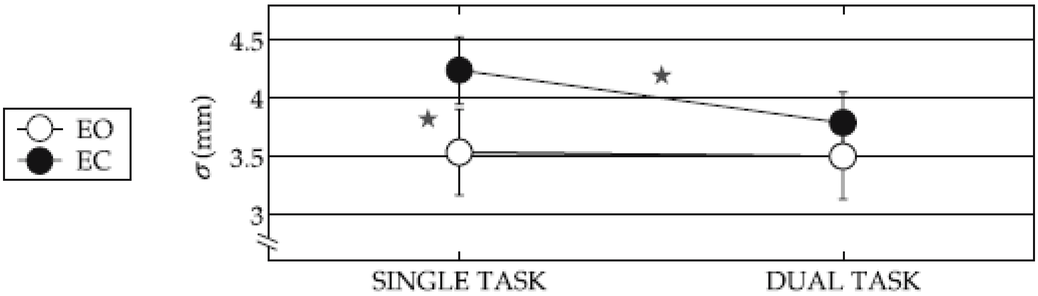


Figure 32 - Legend of the results achieved in (10) (left), graph resuming the standard deviation (right)

We checked that, actually, test 1 and test 2 show the same result of the single-task case, with the values reported in Table 10.

	Test1 - OE	Test2 - CE
AP	1.64 mm	2.85 mm
ML	3.33 mm	6.41 mm

Table 10 - Standard deviation of the AP and ML time series, in the OE and CE conditions

The graphs of the sequential standard deviation of the CoP for test 1 and 2 are reported in Figure 33.

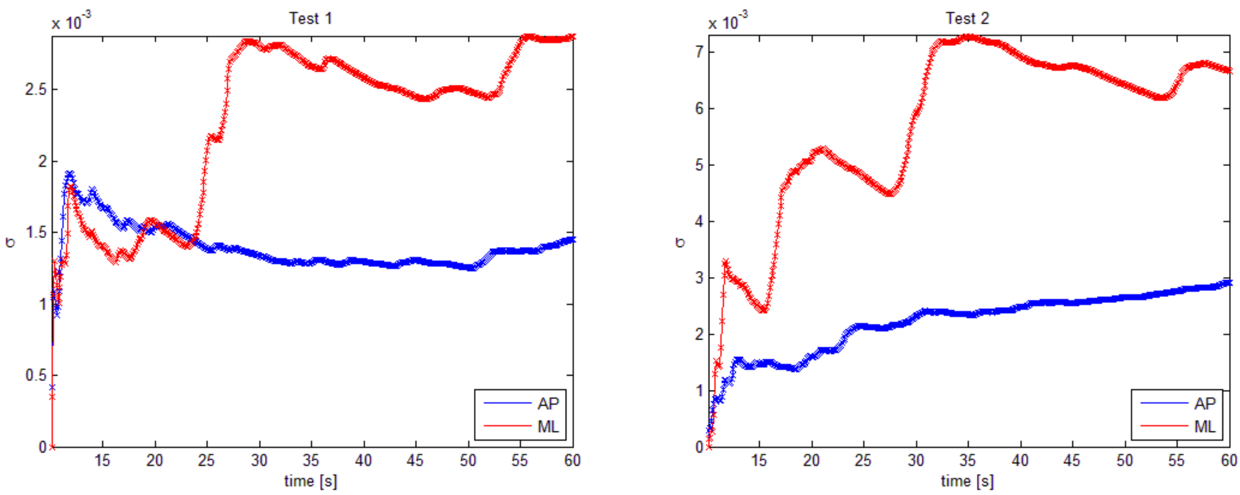


Figure 33 - Sequential standard deviation for test 1 (left) and test 2 (right)

As already known, the CoP is not a stationary signal, and this is proved by the obtained results. The standard deviation for the velocity of the CoP is reported below.

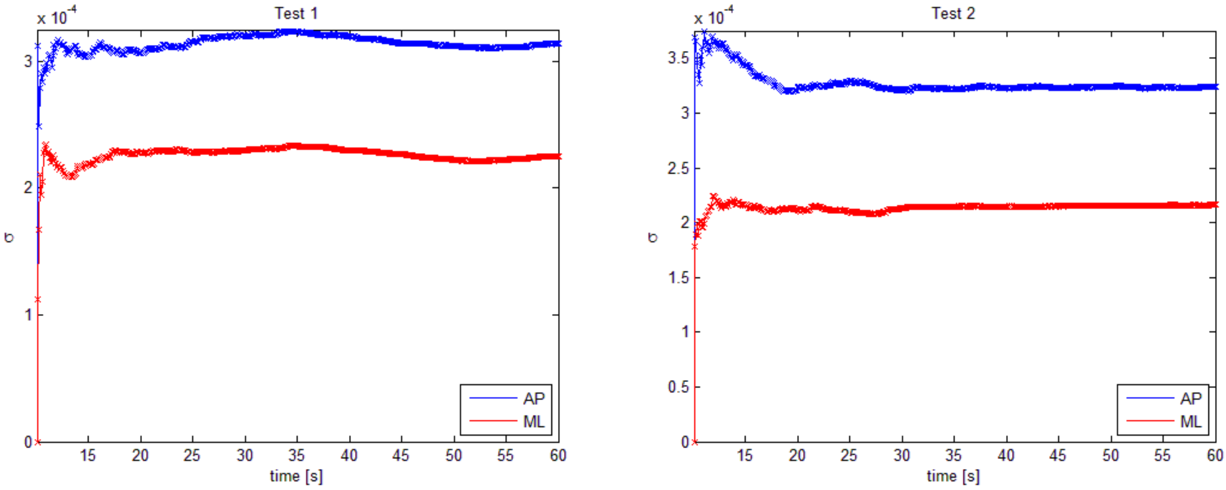


Figure 34 - Sequential standard deviation for the differential time series of test 1 (left) and test 2 (right)

Both AP and ML time series get stable after nearly four seconds. To decide the windows length for the windowed analysis, we used the precedent results. The idea is that, if a parameter is not stable looking at the sequential analysis, it cannot be used for the time-dependent evaluation. In the case of the standard deviation, the CoP displacement is too variable, while the velocity is stable enough after four seconds. Therefore, we studied the velocity of the CoP recurring to a windows length of ten seconds.

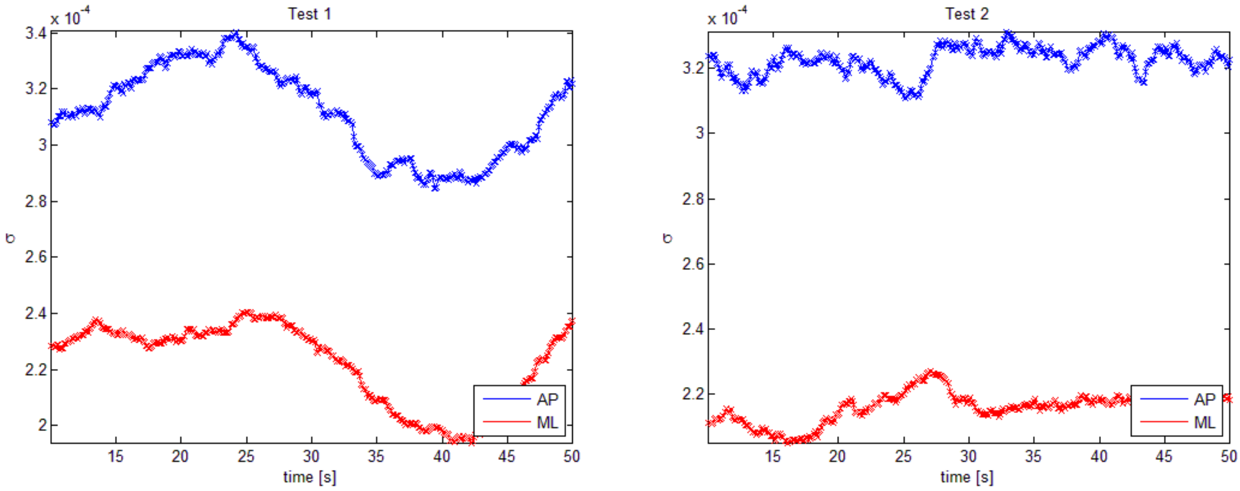


Figure 35 – Test 1(left) and test 2 (right) differential time series, studied with windows of 10 seconds

It can be noticed that there is a fluctuation of the values for Test 1, for both the AP and ML time series. We will discuss this result in the following.

These are the results obtained for test 3. The black background stands for closed-eyes, while the white for open-eyes condition.

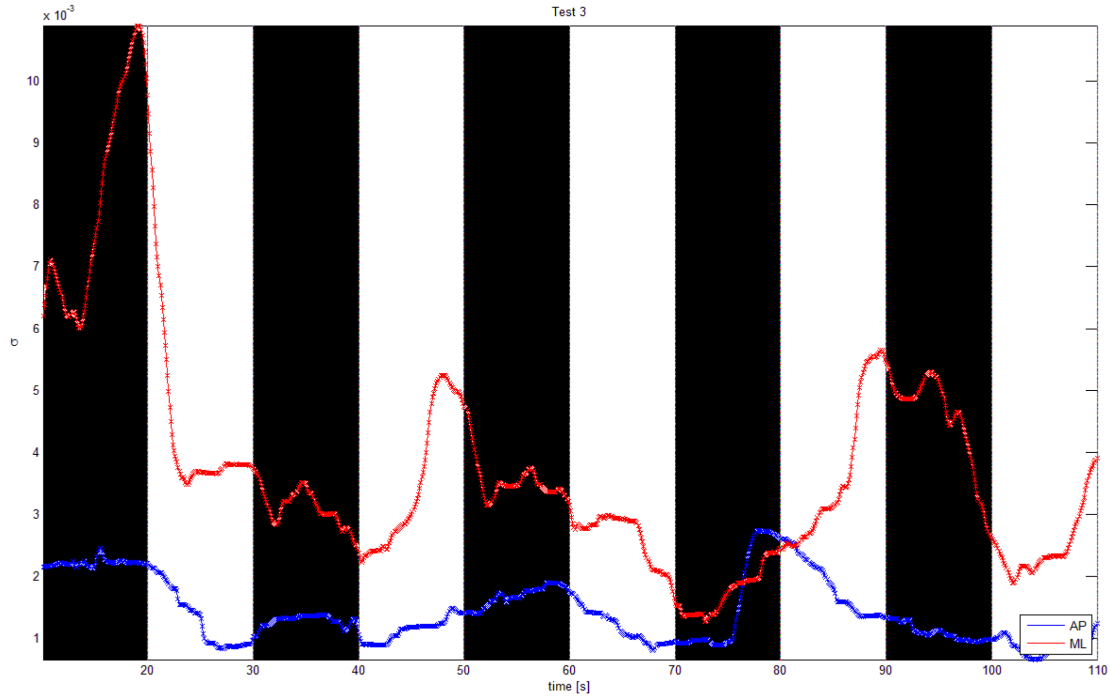


Figure 36 - Standard deviation of the displacement time series of test 3, studied with windows of 10 seconds

We remember that the value at a certain instant t is computed analyzing the interval $[t, t + 10]\ s$. That must be considered when associating a certain condition (given by the background color) to the specific values. Looking at the trends, it seems not possible to recognize whether the subject was in open or closed eyes condition. The same parameters were used for the windowed study of the velocity of the CoP (Figure 37).

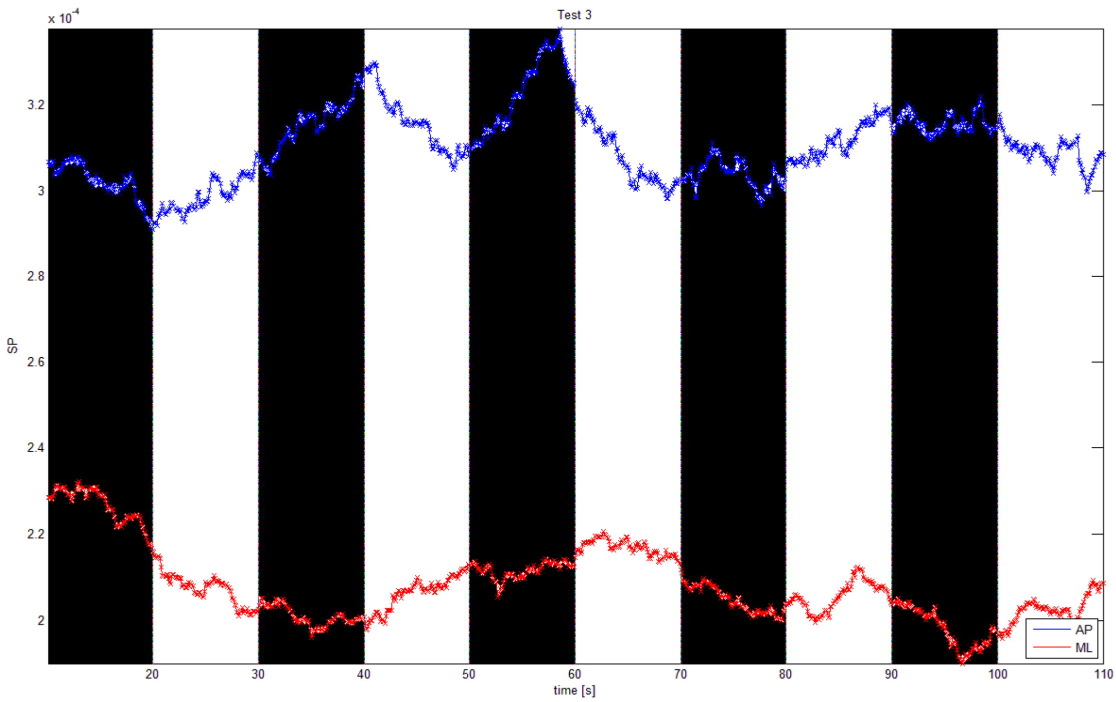


Figure 37 - Differential time series of test 3, studied with windows of 10 seconds

Sway-path length

The results of (10) are reported in Figure 38. In their case, it is not a parameter sensitive to the visual condition.

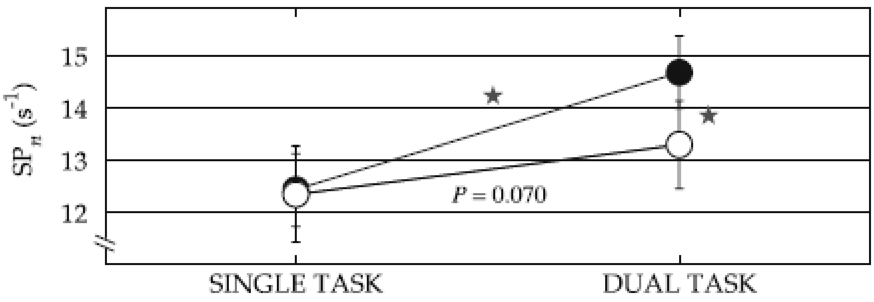


Figure 38 - Graph resuming the sway-path length

In our case, the results in Table 11 show a higher sensitivity.

	Test1 - OE	Test2 - CE
AP	1.6 s^{-1}	0.8 s^{-1}
ML	0.6 s^{-1}	0.2 s^{-1}

Table 11 - Sway-path length of the AP and ML time series, in the OE and CE conditions

The subject did not present similar values in the OE and CE conditions. Furthermore, our values are more distant than five standard deviations from the ones of Figure 38. This could be due to a different normalization factor. Indeed, in (10) the exact formula to calculate SP was not reported, thus we used the one reported in (5).

In the following are presented the results of the analysis for the displacement and the velocity time series (Figure 39).

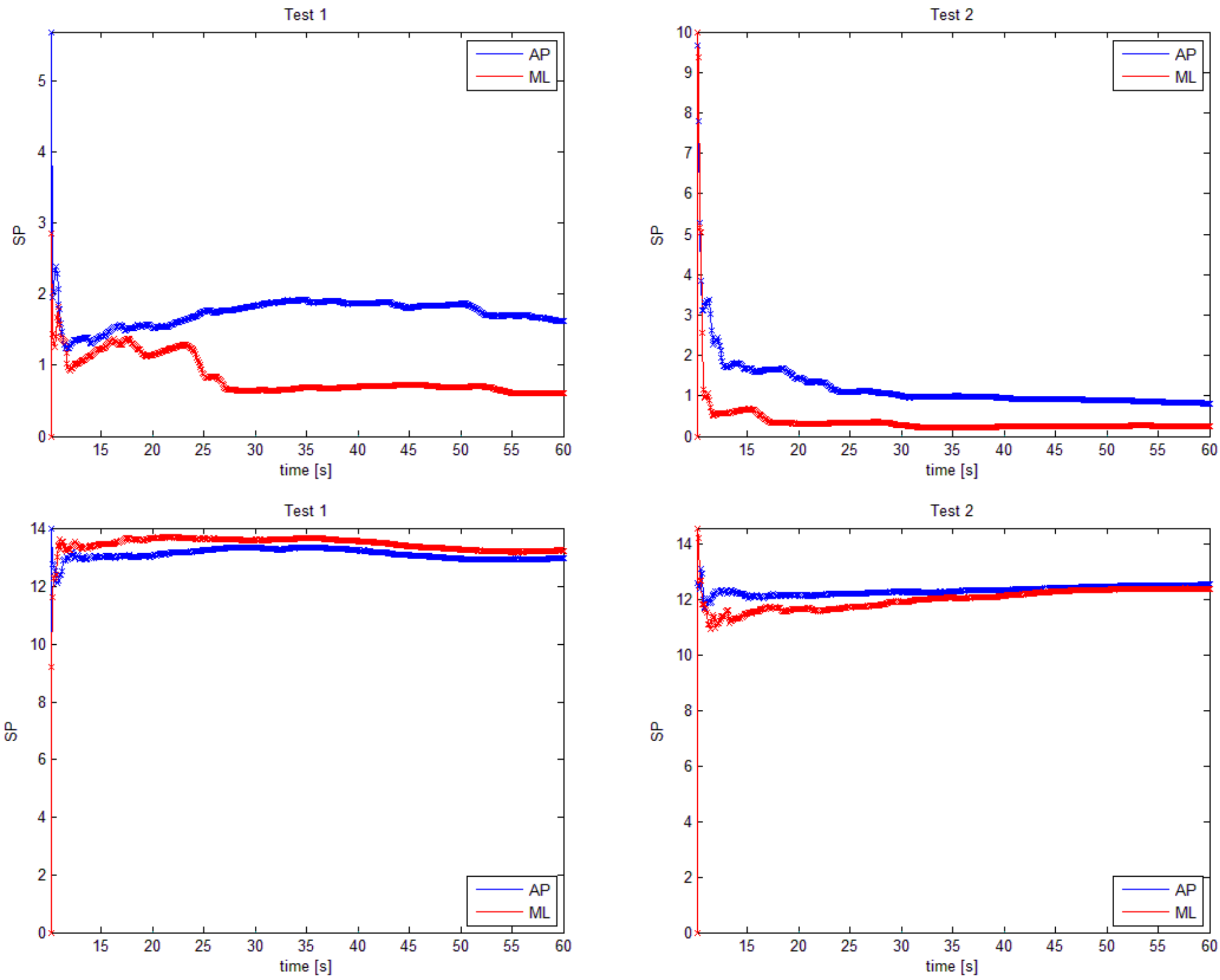


Figure 39 - Sway-path length sequential analysis of the displacement (top) and velocity (bottom) time series. Test 1 is presented on the left, test 2 on the right

Both the velocity and the displacement time series seem more or less stable, especially after ten seconds. Because a higher value of window could not be used (the visual condition changed every 10 seconds), we selected that windows length. In Figure 40 are reported the results for the windowed analyses.

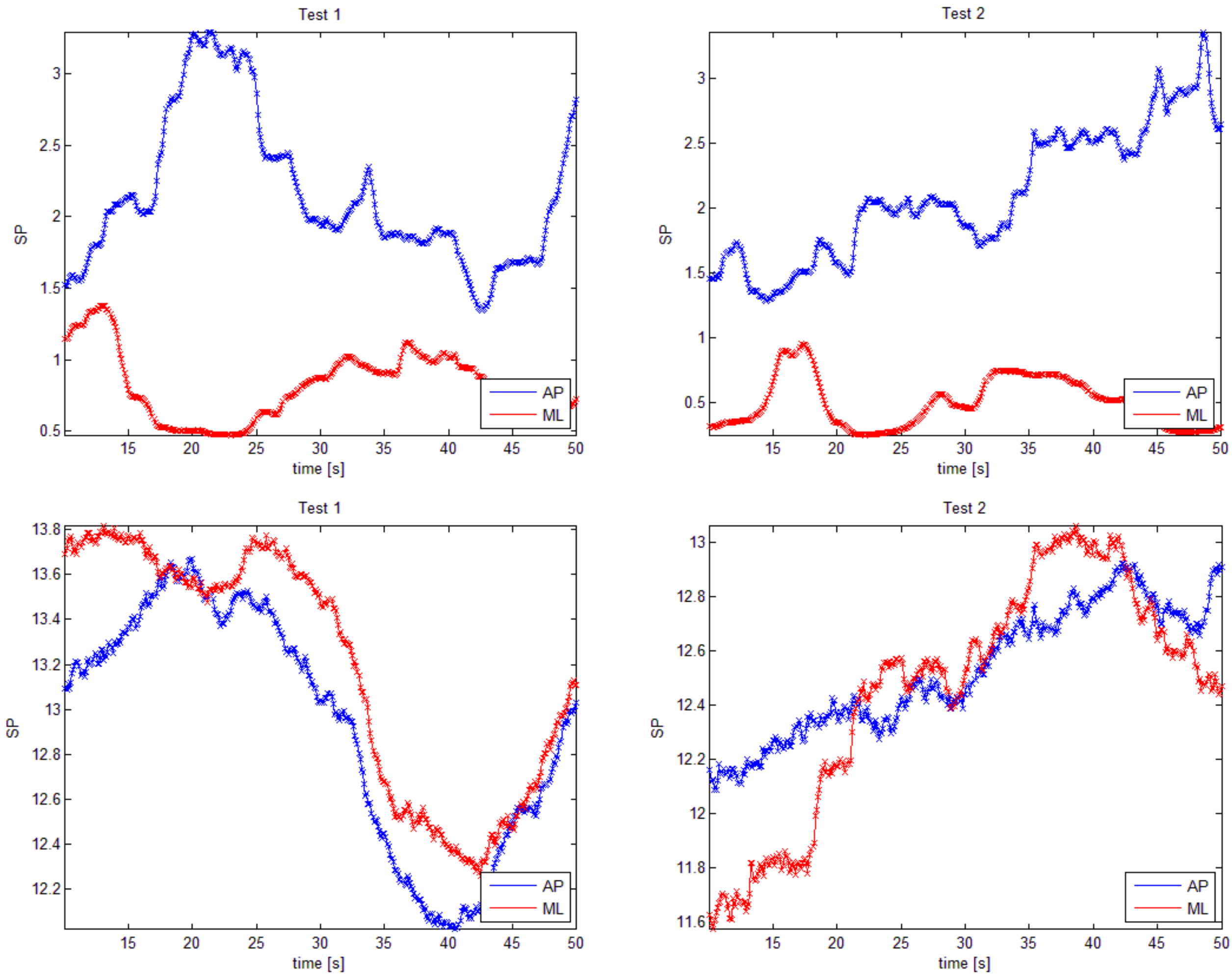


Figure 40 - Sway-path length windowed analysis of the displacement (top) and velocity (bottom) time series. Test 1 is presented on the left, test 2 on the right

In the displacement time series the trends present oscillations that are much clearer for the velocity. The period in the latter case seems the same one obtained for the standard deviation of the velocity time series (Figure 35). A peculiar characteristic related to the visual condition is that the SP of the velocity in the case OE follows a negative trend, while in the CE case follows a positive trend.

Below are reported the result for the windowed analysis of test 3, for both the displacement (Figure 41) and the velocity (Figure 42).

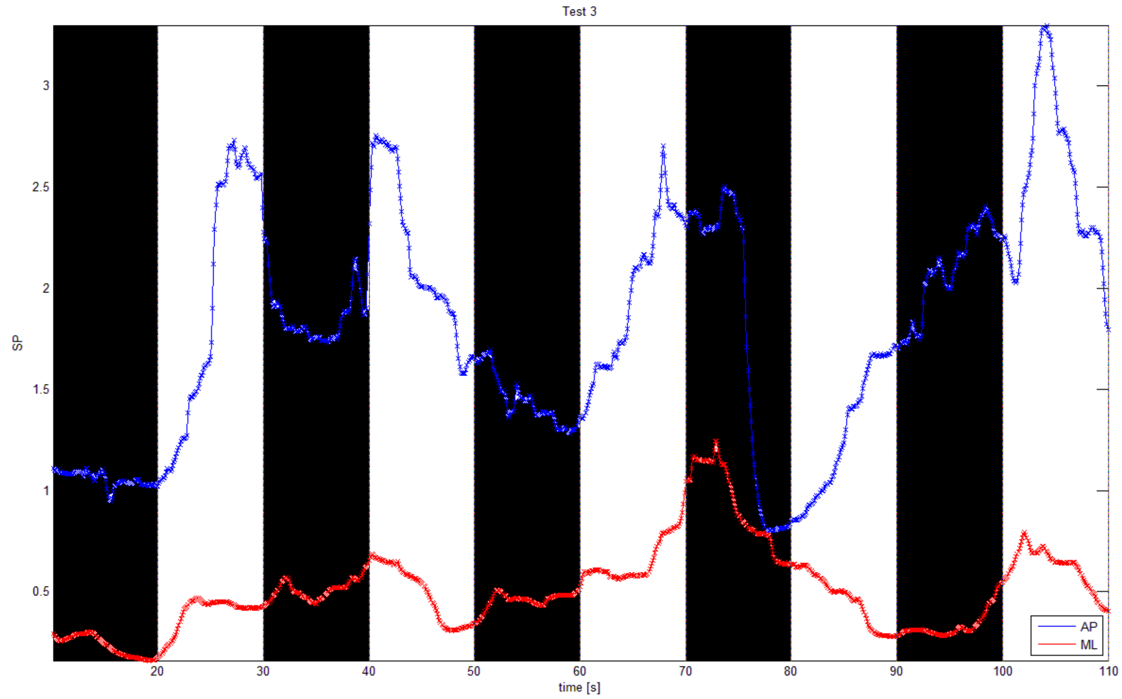


Figure 41 - Sway-path length of the displacement time series of test 3, studied with windows of 10 seconds

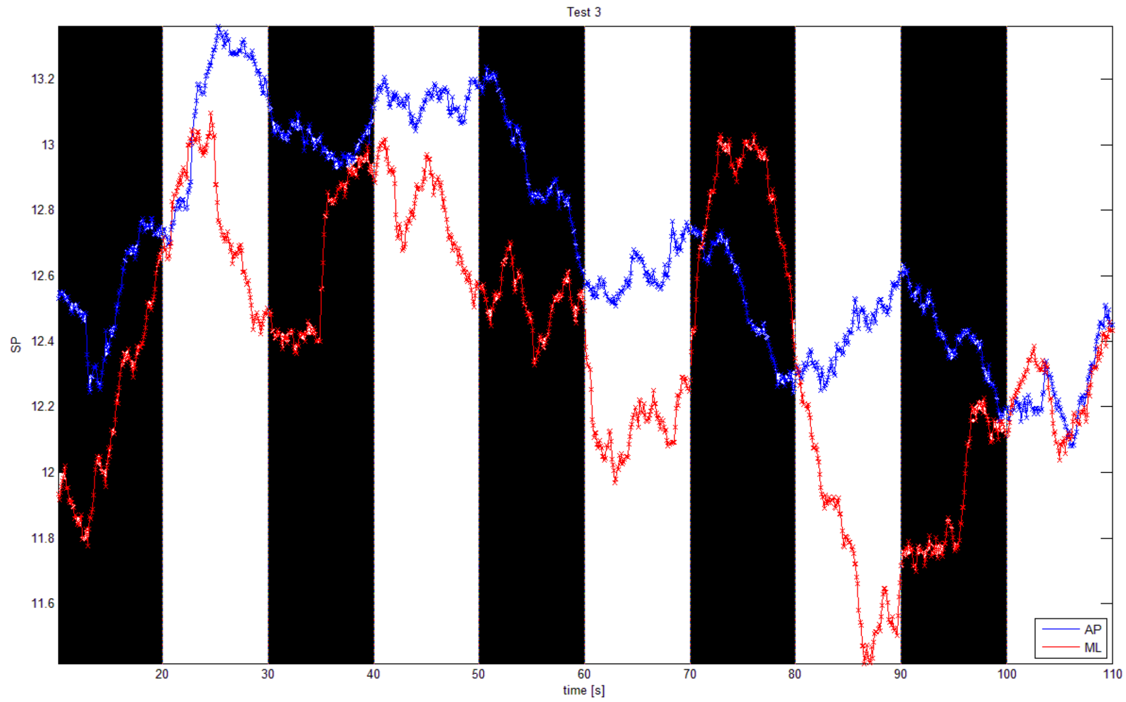


Figure 42 - Sway-path length of the velocity time series of test 3, studied with windows of 10 seconds

It can be noticed that during the CE phase (black column) the SP rises, while in the OE phases it decreases. This is exactly the same trend found in Figure 40, analyzing separately test 1 and test 2.

Representing the slope of the regression line within each interval of 10 seconds, and rescaling the values, it is possible to achieve an oscillatory behavior.

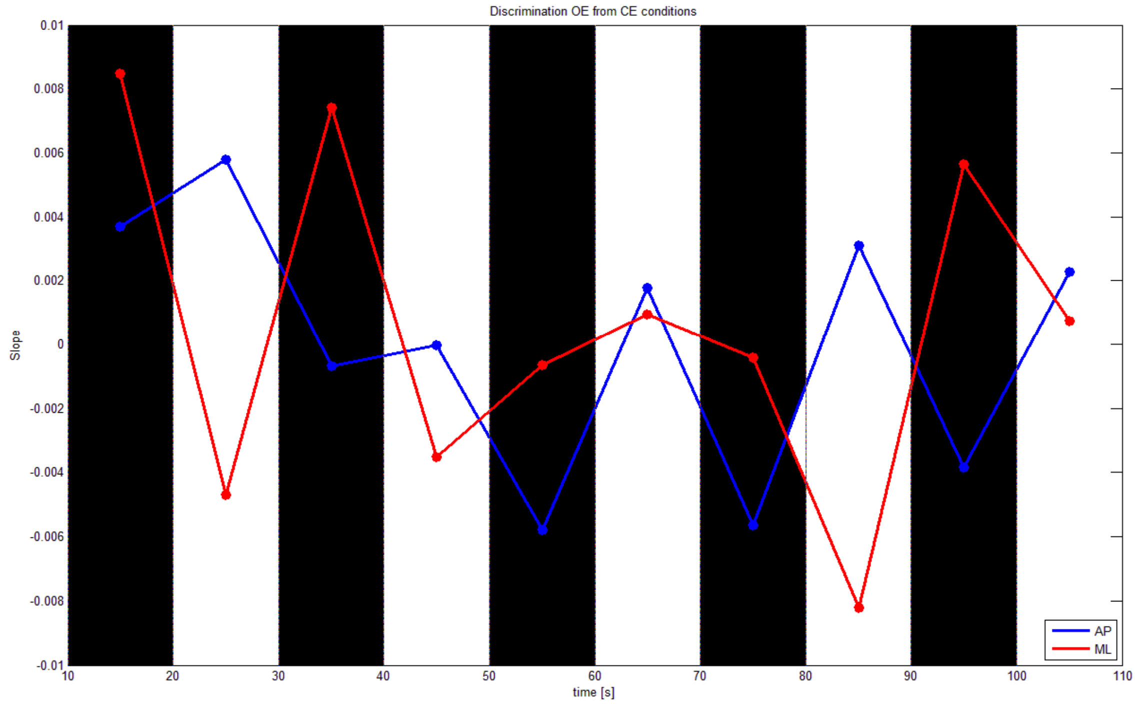


Figure 43 - Slope values of the regression line calculated within each interval of 10 seconds of the SP time series obtained from the study of the velocity of the CoP

From the presented result, it seems sensible to argue that SP is a sensitive parameter for the dynamical study of the visual condition. In this case, the AP time series seemed more reliable compared to the ML time series, that missed the proper trend in the interval from 60 to 70 seconds. Anyway, that could be due to the simplicity of the calculation method or simply the physical characteristics of the subject.

Sample Entropy

As already found also in our case, the SE (for us FE) is a sensitive parameter to the visual condition.

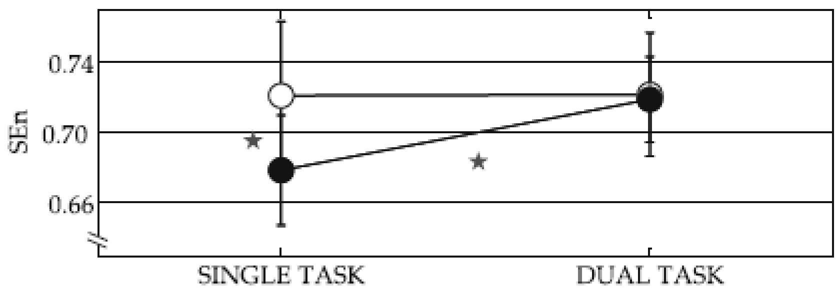


Figure 44 - Graph resuming the sample entropy

Using the set of parameters $m = 4$, $\tau = 0.5\ s$ and $r = 0.5$, the following values were calculated for test 1 and 2.

	Test1 - OE	Test2 - CE
AP	0.53	0.30
ML	0.43	0.31

Table 12 - Fuzzy Entropy of the AP and ML time series, in the OE and CE conditions

As reported in both the results, the entropy is a sensitive parameter to the visual condition. It appears again that the results are strictly dependent on the characteristics of the subjects. Indeed, the precedent dataset outlined exactly the opposite result.

The sequential analyses are showed in Figure 45 and Figure 46.

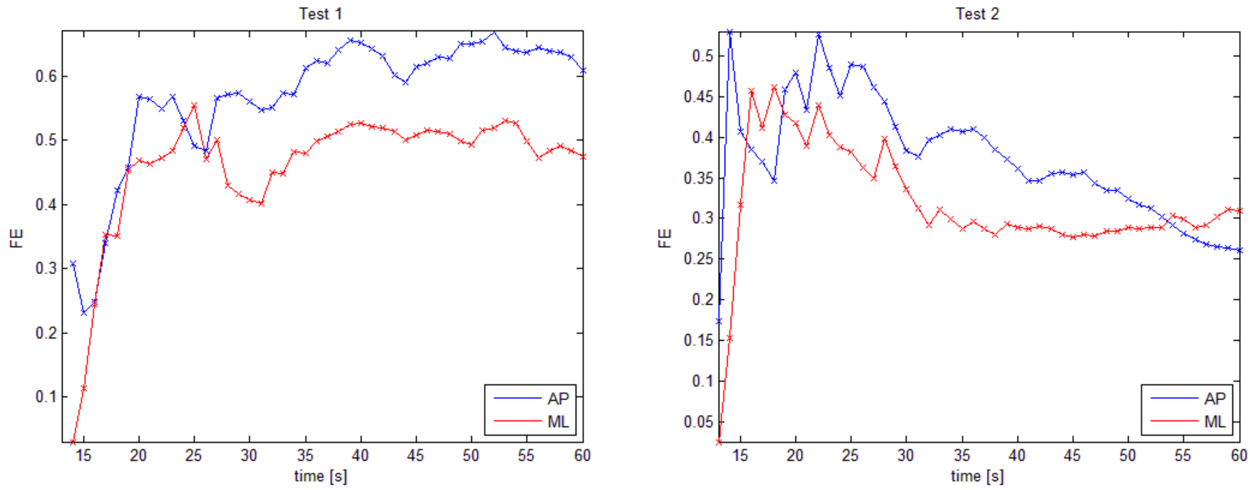


Figure 45 - Sequential Fuzzy Entropy of test 1 (left) and test 2 (right), for both displacement AP and ML time series

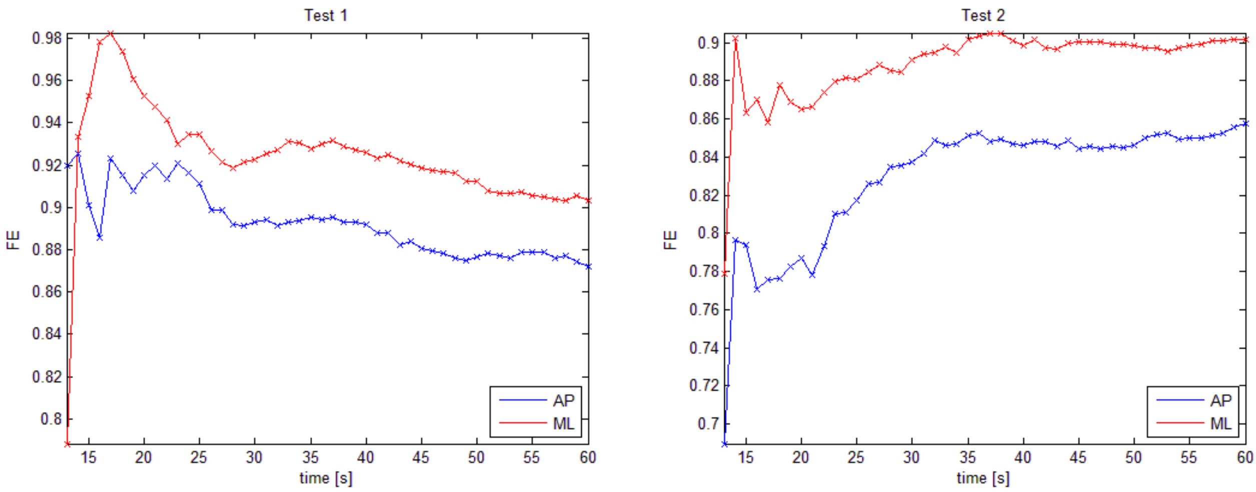


Figure 46 - Sequential Fuzzy Entropy of test 1 (left) and test 2 (right), for both differential AP and ML time series

All the four trajectories are non-stationary. This is particularly true for the displacement time series, while the velocity ones, even if changing with time, maintain a similar trend.

We extended the study to test 3, reporting the results in Figure 47 and Figure 48.

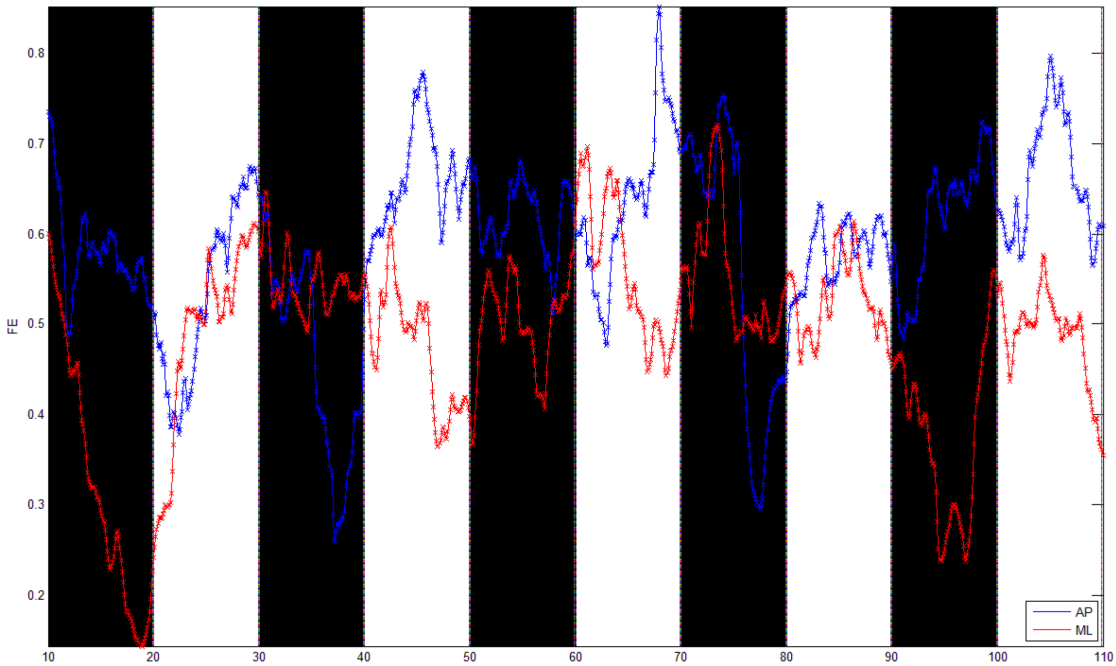


Figure 47 - Fuzzy Entropy of the displacement time series of test 3, studied with windows of 10 seconds

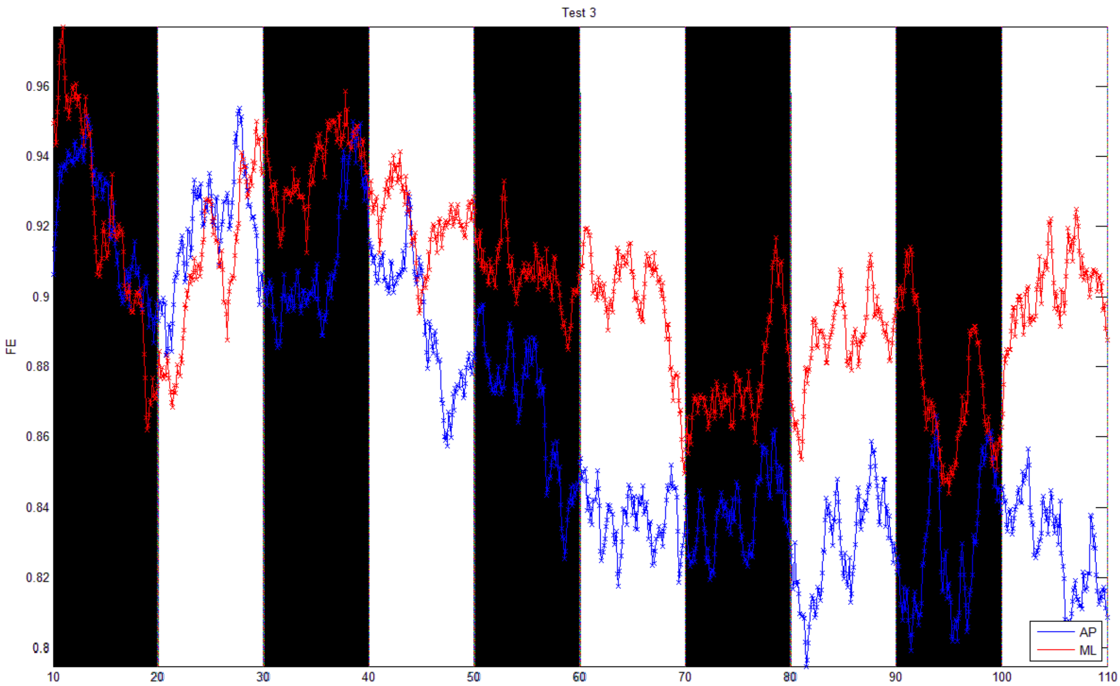


Figure 48 - Fuzzy Entropy of the velocity time series of test 3, studied with windows of 10 seconds

Both the displacement curves present an oscillatory behavior that could seem dependent to the visual condition for certain time intervals, but appear incoherent in others. Roughly speaking, it seems that during the closed-eyes phase, the entropy lowers, while in the opened-eyes phase it rises. That would be coherent with the results obtained analysing separately OE and CE conditions (i.e. what resumed in Table 12).

Hurst exponent

We followed the procedure presented in (34) for the estimation of the Hurst exponent (resumed in 8.0).

The Power Spectral Density (PSD) of test 1 AP time series (Figure 49) outlined that frequencies higher than 8-10 Hz are principally due to noise. Exactly the same trend was found for the ML time series and for tests 2.

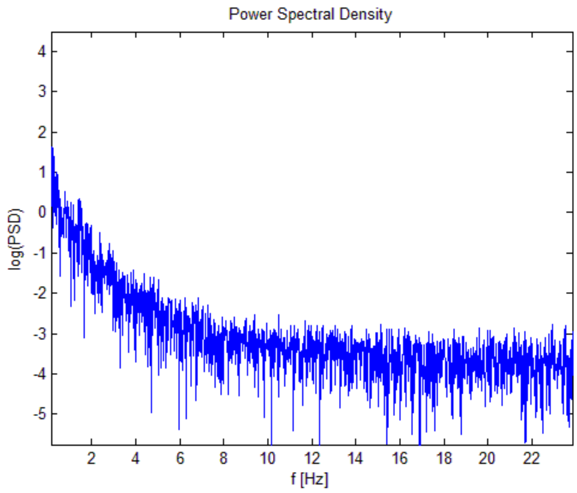


Figure 49 - PSD of test 1 AP time series

In Table 13 the Hurst exponents calculated through lowPSD were reported.

	Test1 - OE	Test2 – CE
AP	0.67	0.64
ML	0.91	0.98

Table 13 - Hurst exponent through lowPSDwe of the AP and ML time series, in the OE and CE conditions

Because the results outlined the fractional Brownian motion nature of the CoP, and we had $0.2 < H < 1$, the best algorithm for the evaluation of the Hurst exponent was the Scaled Windowed Variance (SWV). In particular, we used the linear detrended Scaled Windowed Variance (ldSWV), an improved version of SWV. As showed in (34), and reported in Figure 50, this algorithm is really accurate in the estimation.

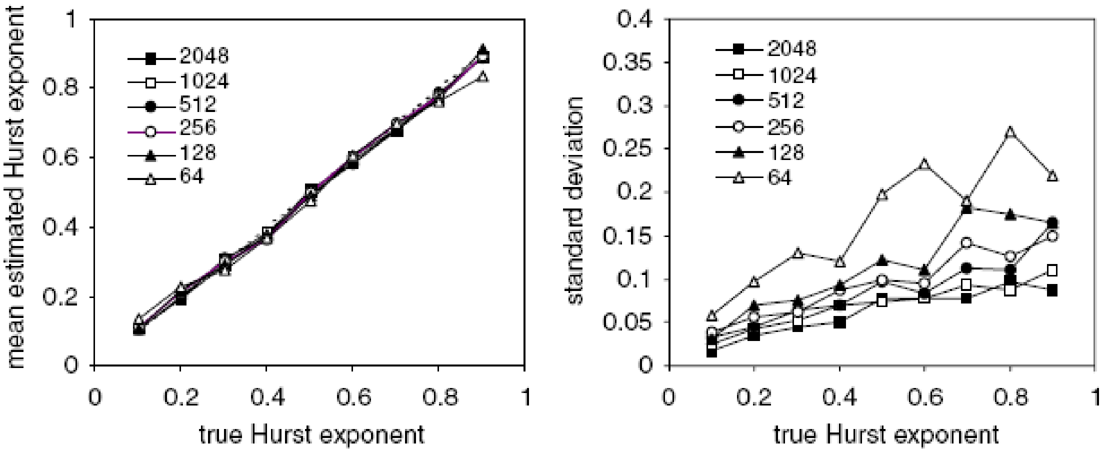


Figure 50 - SWV analysis. Plots of mean estimated Hurst exponent versus the true H (left), and estimated H standard deviation versus true H (right). The different symbols refer to the length of the used time series

Because of the high number of samples of our time series (> 2048), it is reasonable to expect that the Hurst exponent estimations are reliable (for the sensitivity to noise of SWV, see the details in (34)).

In the following tables are resumed the results of (10) (Table 14) and ours (Table 15).

	Condition	Mean	Vision (EO vs. EC)			Condition	Mean	Dual task (ST vs. DT)		
			$F(1, 29)$	P	f			$F(1, 29)$	P	f
α	EO	1.39	13.70	<0.001	0.69	ST	1.39	24.57	<0.001	0.92
	EC	1.34				DT	1.35			

Table 14 - Scaling exponent results

	Test1 - OE	Test2 - CE
AP	0.26	0.21
ML	0.49	0.52

Table 15 - Hurst exponent through SWV of the AP and ML time series, in the OE and CE conditions

It is worth to remember the relation between the scaling exponent α obtained through DFA, reported in Table 14, and the Hurst exponent:

$$H = \alpha - 1$$

Therefore, in both the cases the Hurst exponent is higher in the OE condition. Moreover in both the cases the Hurst exponent is in the persistent range.

Now studying the sequential time series:

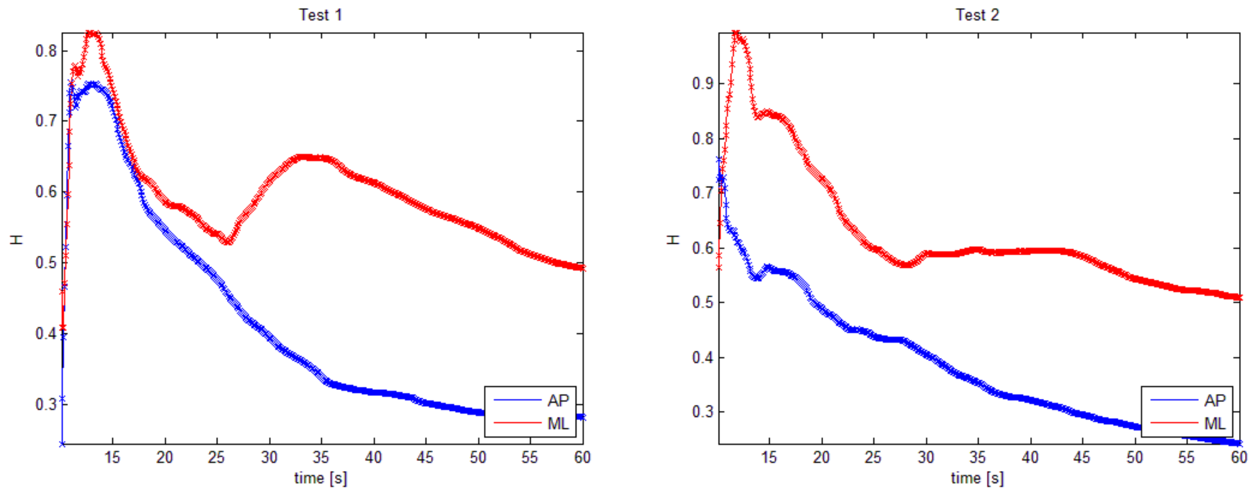


Figure 51 - Sequential Hurst exponent of test 1 (left) and test 2 (right), for both displacement AP and ML time series

Within the first ten seconds the AP time series is in the persistent range, while after it is in the anti-persistent range. This is true for both OE and CE time series. The true difference, in this case, stands between AP and ML time series. Indeed, the ML time series is almost always in the persistent range.

In the following the results for test 3.

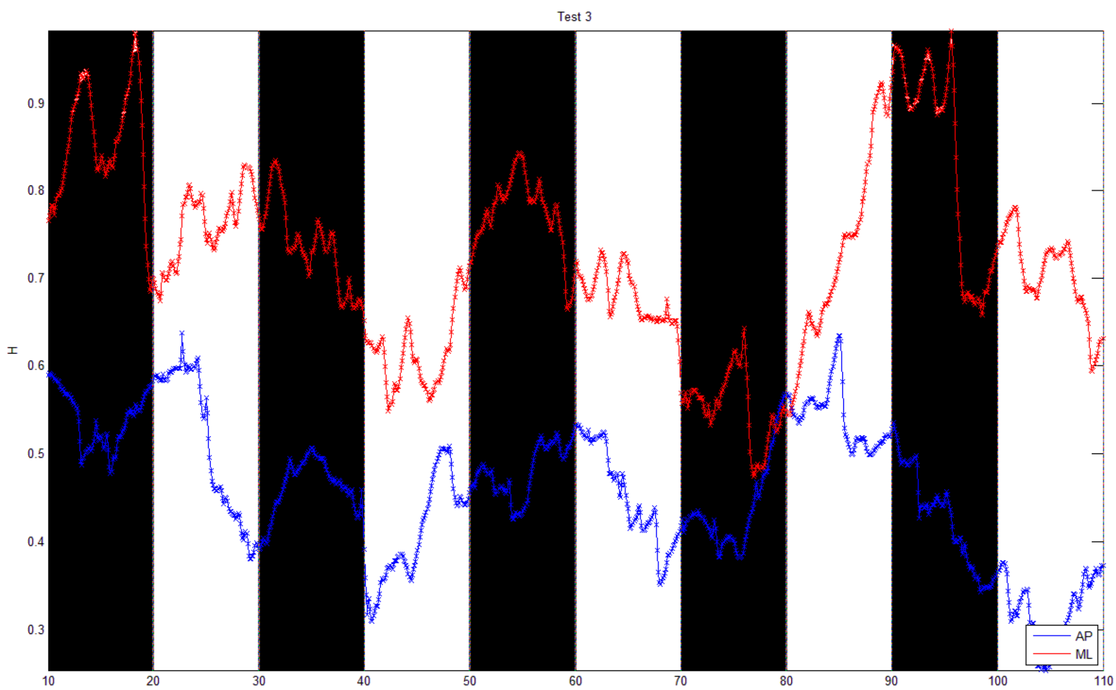


Figure 52 - Hurst exponent of the displacement time series of test 3, studied with windows of 10 seconds

The Hurst exponent of the velocity time series was not studied because it is the same of the displacement time series (34). That is due to the mirror effect, introduced in chapter 2. The differences can only be due to a numerical inaccuracy from one or both the estimations.

7.3 Dynamical study of the cognitive condition

Here are presented the results for test 4 of the specific dataset. In test 4 the single-task and dual-task conditions were alternated, recording the time series with eyes closed.

For problems related to a data transfer, the other test could not be studied.

From now on, we will depict the graphs using three colors:

- Dual-task condition, spelling the inverted word
- Dual-task condition, listening to word, trying to invert it
- Single-task condition

Each time series was studied with the windowed approach, analyzing windows of 10 second. The length was chosen taking as a reference the study of the incremental time series, proposed in the precedent paragraph.

Our results were confronted, as already made for the visual condition, with the ones obtained in (10), reported in Figure 53 and .

Table 2 Main and interaction effects| of vision and dual task (i.e., collapsed over x and y time-series) of sample entropy (SEn), standard deviation (σ), sway-path length of the normalized (by the standard deviation) posturogram (SP_n), largest Lyapunov exponent (λ_{max}) and scaling exponent (α) of COP time-series for 30 healthy individuals

	Condition	Mean	Vision (EO vs. EC)			Condition	Mean	Dual task (ST vs. DT)			Vision × dual task[a]		
			$F(1, 29)$	P	f			$F(1, 29)$	P	f	$F(1, 29)$	P	f
SEn	EO	0.72	3.83	=0.060[*]	0.36	ST	0.70	1.45	ns	0.25	6.72	<0.05	0.48
	EC	0.70				DT	0.72						
σ	EO	3.52	11.82	<0.005	0.64	ST	3.89	2.45	ns	0.29	3.18	=0.085	0.33
	EC	4.01				DT	3.64						
SP_n	EO	4.27	5.28	<0.05	0.43	ST	4.13	13.57	<0.005	0.68	6.98	<0.05	0.49
	EC	4.52				DT	4.66						
λ_{max}	EO	1.56	36.23	<0.001	1.12	ST	1.71	0.10	ns	0.06	4.26	<0.05	0.38
	EC	1.88				DT	1.73						
D_2	EO	2.23	23.58	<0.001	0.90	ST	2.20	45.70	<0.001	1.26	6.15	<0.05	0.46
	EC	2.48				DT	2.51						
α	EO	1.39	13.70	<0.001	0.69	ST	1.39	24.57	<0.001	0.92	1.80	ns	0.25
	EC	1.34				DT	1.35						

[*] Significant vision × plane interaction ($F(1, 29) = 5.48$, $P < 0.05$, $f = 0.44$), which was caused by the fact that, in contrast to that in the frontal plane, the effect of vision was significant in the sagittal plane ($F(1, 29) = 6.47$, $P < 0.05$, $f = 0.47$)

[a] See Fig. 2 for mean values of the conditions EO-ST, EC-ST, EO-DT and EC-DT

Figure 53 - Report of the results achieved in (10)

We recall that the scaling exponent α obtained through Detrended Fluctuation Analysis, used in (10), it is linked with the Hurst exponent we calculated through Scaled Windowed Variance, by the relation:

$$H = \alpha - 1$$

Standard deviation

Beginning with the standard deviation, the results for both displacement and velocity are depicted respectively in Figure 54 and Figure 55.

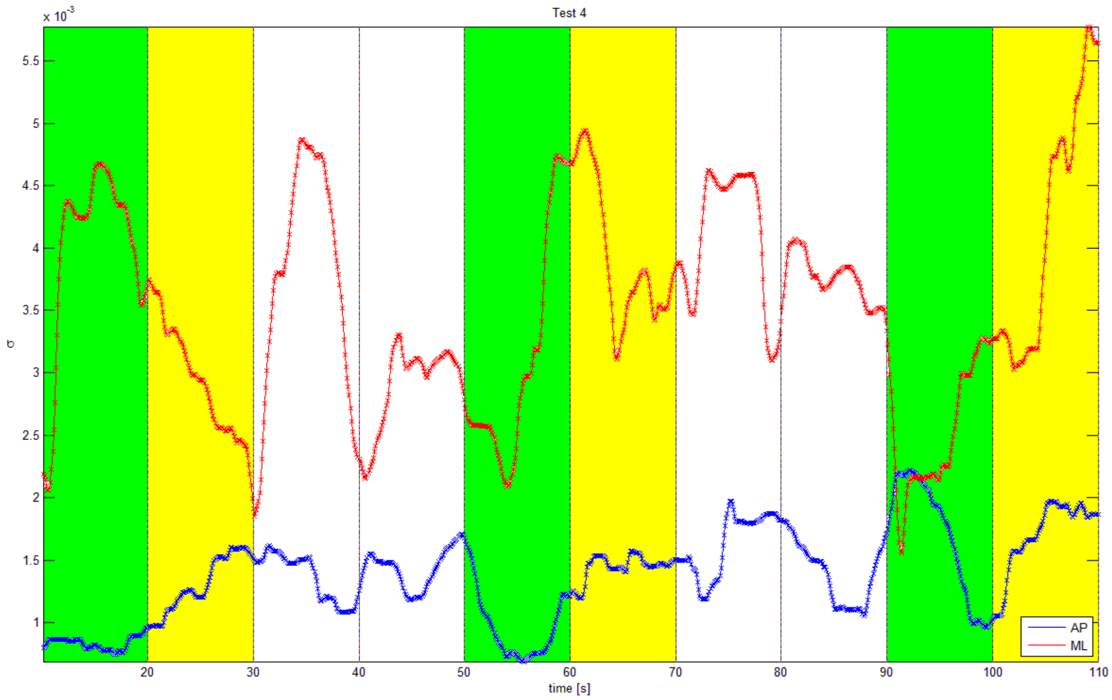


Figure 54 - Standard deviation of the displacement time series of test 4, studied with windows of 10 seconds

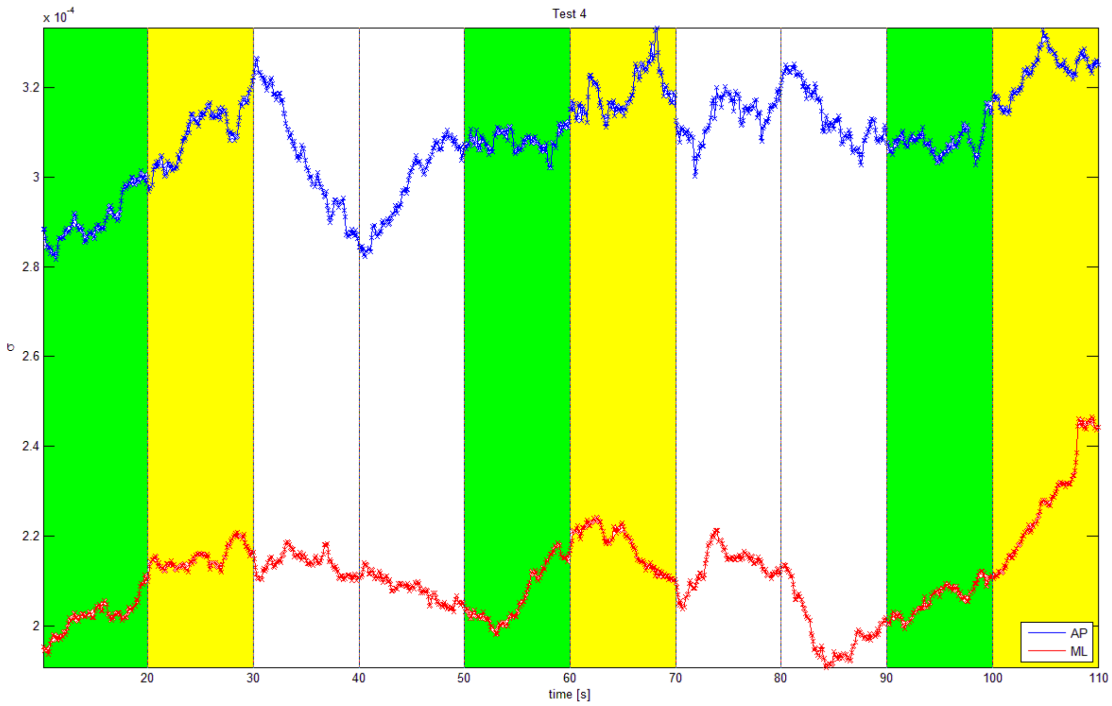


Figure 55 - Standard deviation of the velocity time series of test 4, studied with windows of 10 seconds

The trends look similar to the ones proposed in the study of the visual condition. Especially in the displacement time series, it can be seen an oscillatory behavior hardly due to the cognitive condition. Furthermore, in (10) the standard deviation during the dual-task is lower than the standard deviation during the single task, in the closed eyes condition (see Figure 32), but we do not notice a similar behavior (it seems completely uncorrelated to the cognitive condition).

Sway-path length

Below the same study was extended to the sway-path length. Only the AP time series of the displacement time series seems somehow correlated with the dual task. In particular, it seems that while listening to the word and trying to invert it (green background) the value increases. Then, depending on the amount of time involved to spell out the name, the SP stays high until there is a sudden drop (yellow background).

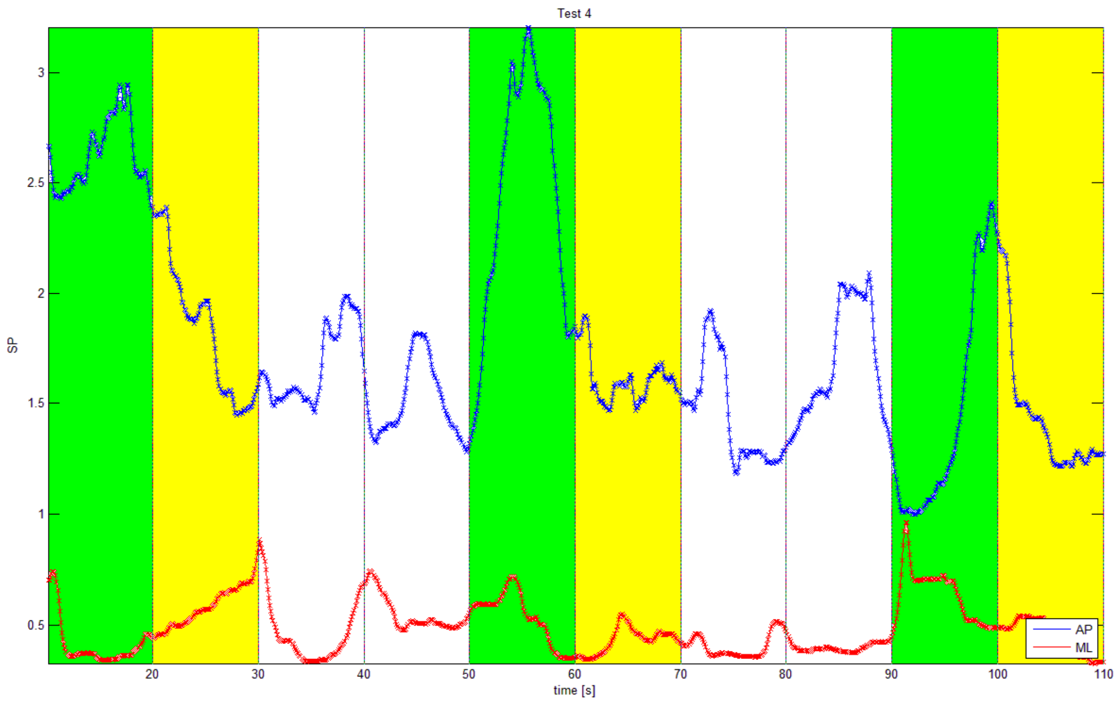


Figure 56 - Sway-path length of the displacement time series of test 4, studied with windows of 10 seconds

In (10) the SP in the dual-task is higher than the one in the single-task (Figure 38). Their result is compatible with ours.

Nothing in particular arises from the analyses of the velocity (Figure 57). We recall that the SP analysis of the velocity was the only one which gave a good result in the dynamic recognition of the visual condition (Figure 43).

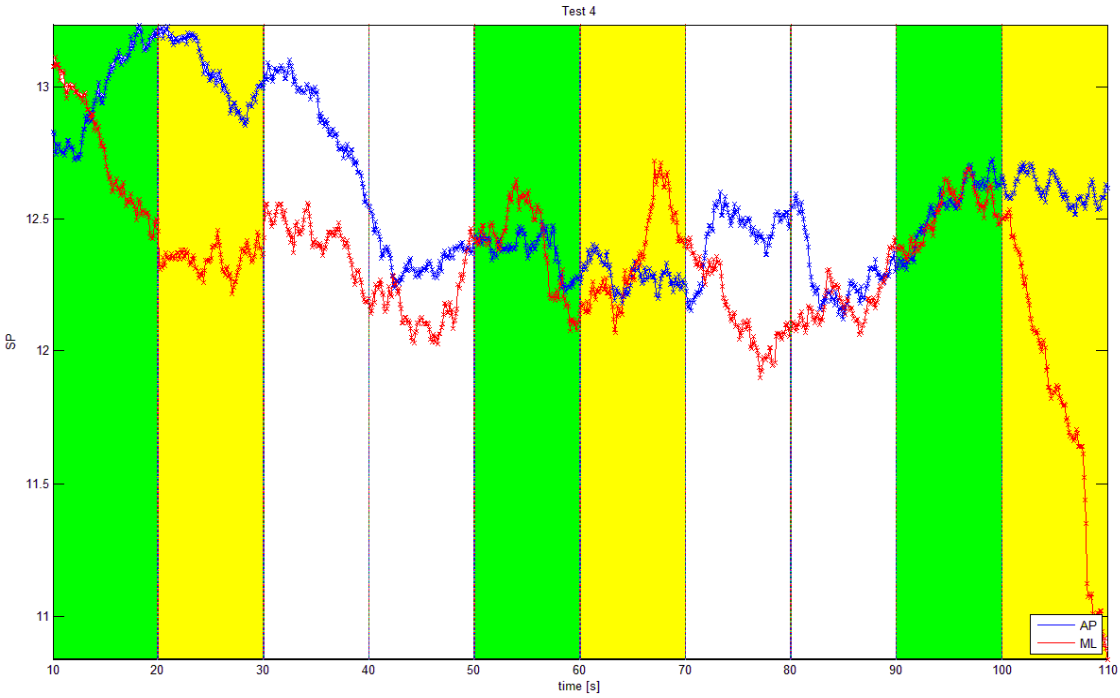


Figure 57 - Sway-path length of the velocity time series of test 4, studied with windows of 10 seconds

Fuzzy Entropy

As already seen for the visual condition, the Fuzzy Entropy appears as a highly variable signal. The qualitative analysis of the obtained results does not allow any particular consideration.

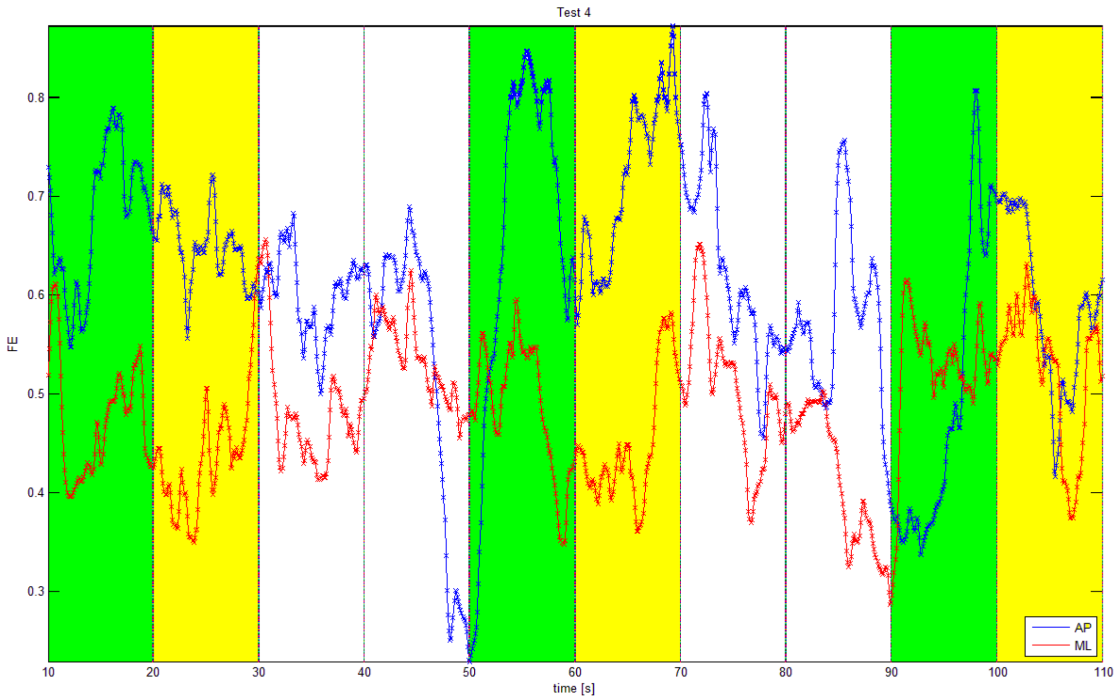


Figure 58 - Fuzzy Entropy of the displacement time series of test 4, studied with windows of 10 seconds

In (10) the Sample Entropy gets higher when the subjects are in the dual-task condition. Our results seem to point in the opposite direction, if we look at the trends in the green and yellow backgrounds. However it seems also the contrary if we consider that the value at the beginning of the green background is the one that better describes the mental activity due to the inversion, because it was computed considering the first 10 seconds of mental activity, and two values out of three (60 seconds and 90 seconds) have the lowest values of entropy. That is mostly true for the AP time series.

Hurst exponent

The analysis of the Hurst exponent was the only one capable to show a form of correlation with the dual-task. In Figure 59 are reported the trends for both AP and ML displacement time series, while in Figure 60 the AP time series alone was represented. To explain more clearly our position, only the values at the beginning of each interval was represented, and plotted in the middle of the intervals.

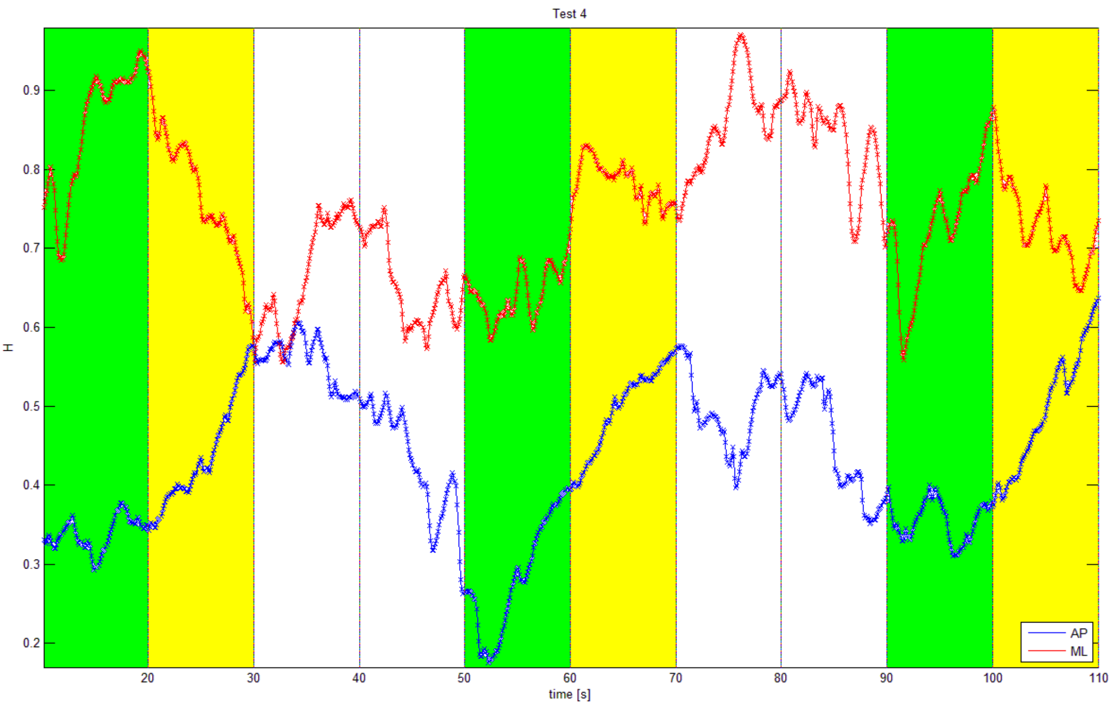


Figure 59 - Hurst exponent of the displacement time series of test 4, studied with windows of 10 seconds

In (10) the values of the scaling exponent lower during the dual-task. That is exactly what happens in our case. Most interestingly, we noticed that during the single-task the Hurst exponent of the AP time series is mainly in the persistent region, while during the dual-task, it is in the antipersistent region. Figure 60 reports that trend, outlining the value $H = 0.5$, that is the boundary between the persistent and the antipersistent region (for details on the meaning of H, see chapter 2).

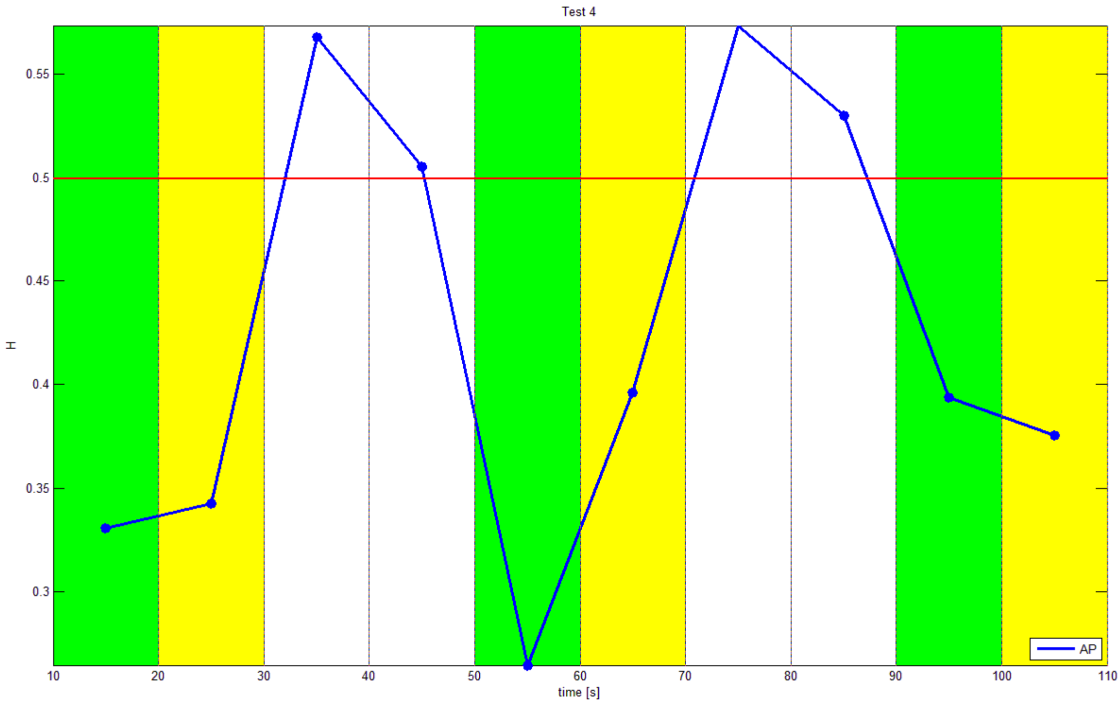


Figure 60 – Hurst exponent values calculated within each interval of 10 seconds of the AP displacement time series

8. Dimensionality of postural sway

The statistical test revealed that the complex signals created joining the AP with the ML time series, respectively as the real and imaginary part of the complex signal, are non-circular. That was true for each trial of each subject in each visual condition. This result is not outstanding, thinking at the definition of circularity. Indeed, circularity can be due to two reasons: being $z = x + jy$ a sample of the complex signal, x and y must have equal variances and x and y must be uncorrelated. Looking at the precedent results showed on this thesis, it is clear that none of these statements are true. Correlation between samples was shown with the complex-nature test, while the differences in the variances was proven with the time-dependent study of the CoP.

Proved the non-circularity of the CoP, we implemented the M-step ahead prediction algorithm recurring to CLMS, ACLMS, QLMS and AQLMS. Because of the non-circularity of the signal, we expected better performances in the case of the augmented version of the adaptive filters. Below we reported the results.

On time series of 5000 samples (the first 10 seconds, i.e. 1000 samples, were neglected as always), we used 3900 samples for the training and 1000 samples for the test. To avoid any possible help in the prediction, we separated training and test time series with 100 not used samples.

Complex case: OE

Starting with the open-eyes case, in Figure 61 are reported the prediction gains for both the training and the test phases. The prediction gain was set to 0.14 seconds (14 samples).

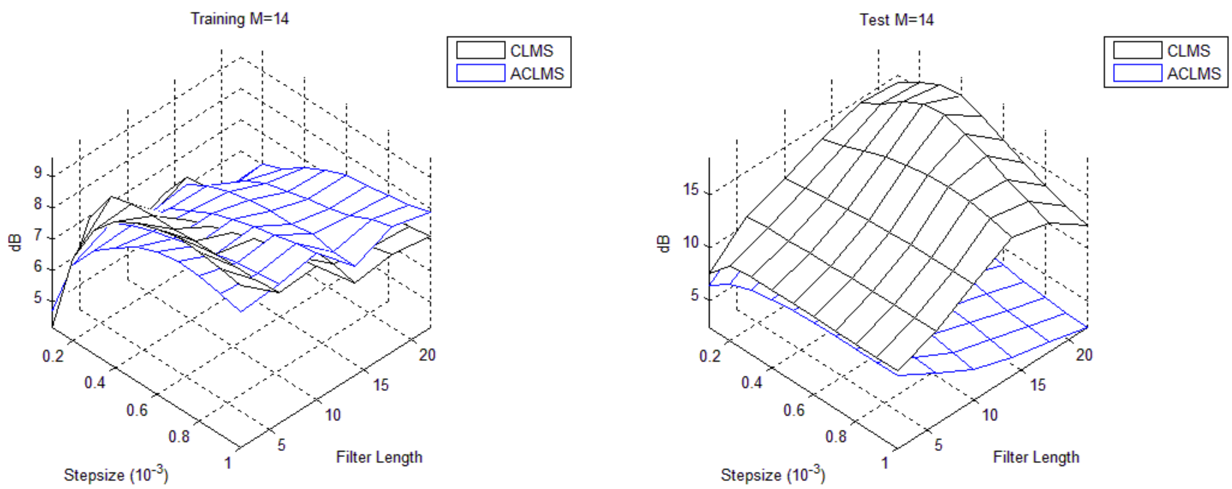


Figure 61 - OE case: CLMS and ACLMS performances as function of the stepsize and filter length, for both the training (left) and test (right) phases. The prediction horizon was set to 14 samples (0.14 seconds)

Because of the prediction horizon, the values of the gain are not so high, especially in the training phase. Starting from the latter, we see that the gain rises with the value of the stepsize and lowers with the order of the filter. The rises in the stepsize allows the filter to follow more quickly the desired signal, while the augment of the order probably causes problems in the convergence of the algorithm. It can be noticed that the ACLMS performs better in this phase, as reasonable because of the non-circular nature of CoP. During the test, unexpectedly, the CLMS performed much better than the ACLMS, especially for filter of high order. For the ACLMS the performamces worsened with the rising of the order of the filter. That could be due to an overfitting of the model arising only in the case of the augmented filter, maybe because of the higher information it handles. Another possibility, looking at the results on the wind velocity elaborated in (17), it is that a higher number of samples for the test it is needed to let the augmented filters converge (usually they are slower than their non-augmented counterparts).

Studying the dependence to the prediction horizon, we achieved the results reported below.

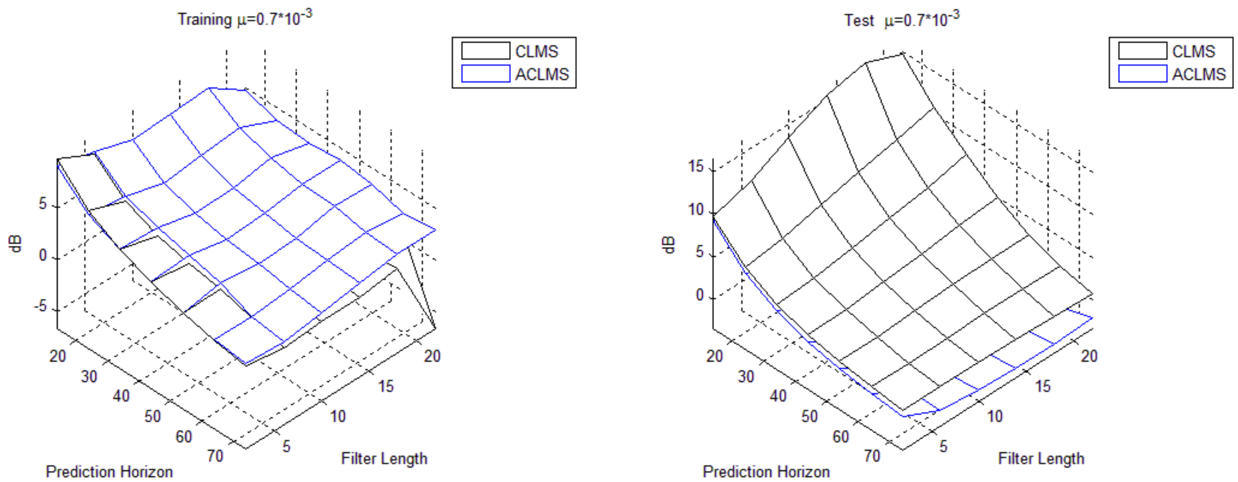


Figure 62 - OE case: CLMS and ACLMS performances as function of the prediction horizon and filter length, for both the training (left) and test (right) phases. The stepsize was set to $0.7 * 10^{-3}$

As reasonable, the prediction gain lowers as the prediction horizon augments. Again the ACLMS performs slightly better than the CLMS during the training, but worst in the test phase. Filters of higher order perform better in the test phase for lower prediction horizon. For the other values, and in the training phase, the performances remain almost the same at every filter length.

Complex case: CE

In Figure 63 and Figure 64 are reported the performances as function of stepsize and filter length, and prediction horizon and filter length, respectively.

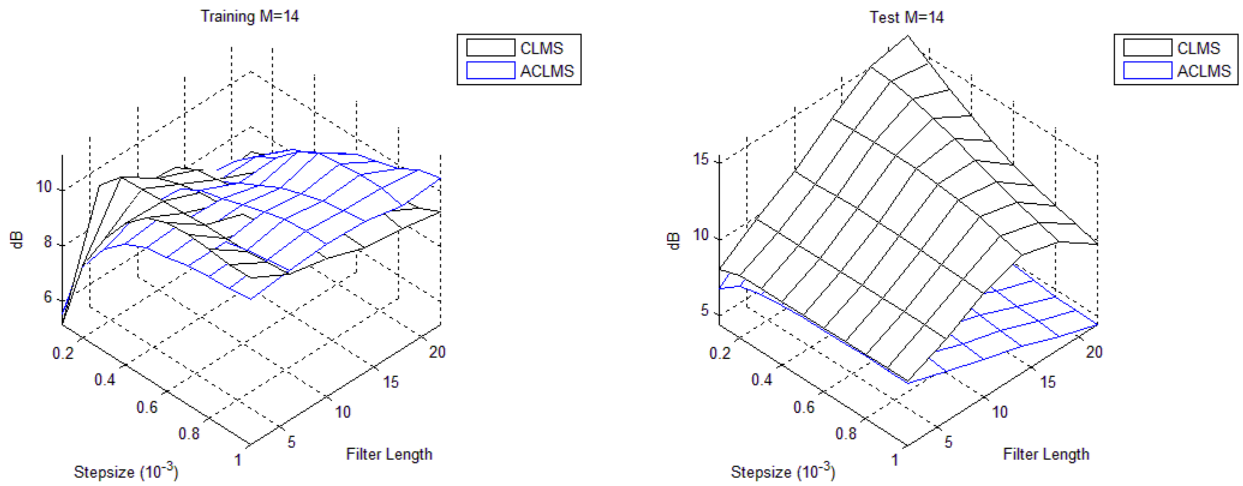


Figure 63 - CE case: CLMS and ACLMS performances as function of the stepsize and filter length, for both the training (left) and test (right) phases. The prediction horizon was set to 14 samples (0.14 seconds)

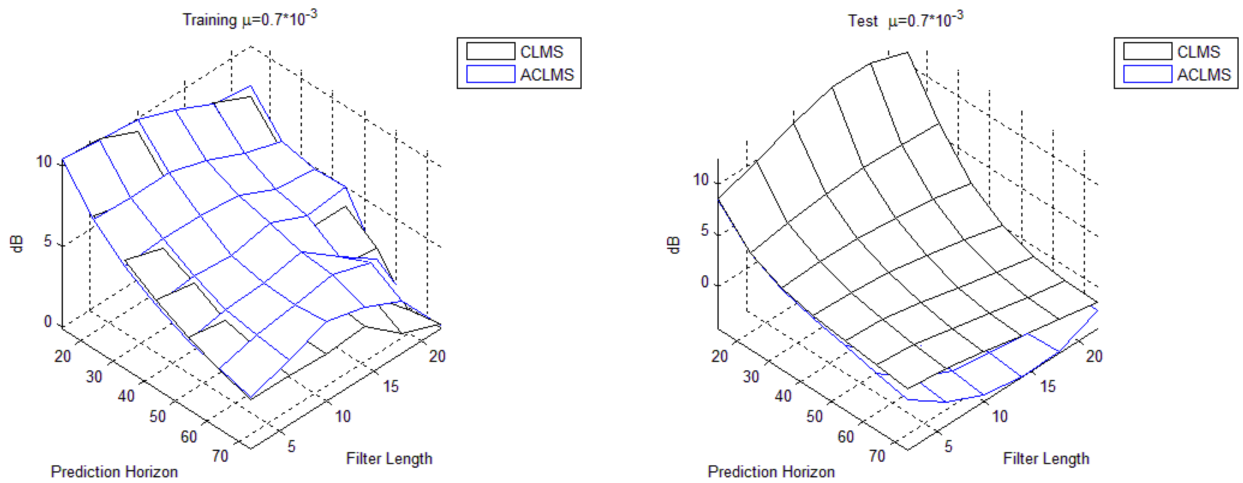


Figure 64 - CE case: CLMS and ACLMS performances as function of the prediction horizon and filter length, for both the training (left) and test (right) phases. The stepsize was set to $0.7 * 10^{-3}$

In both the cases, the considerations are the same presented in the open-eyes case. The only two differences are the values of the gain, generally higher in the CE case.

Quaternion case and confront with the complex case

During the study of the quaternion model, we separated the results in OE and CE, to compare them with the ones achieved in the complex domain. In Figure 65 the results for test and training are reported. Again, during the training phase the augmented filter performs better than the non-augmented, while the contrary is true for the test phase. We find smoother profiles of the gain compared with the ones achieved before, especially for the training. There is a slightly improvement of the performances with the filter length in all the cases but the test for the AQLMS. As a final remark, the closed-eyes case confirms to achieve higher prediction gains, in both training and test.

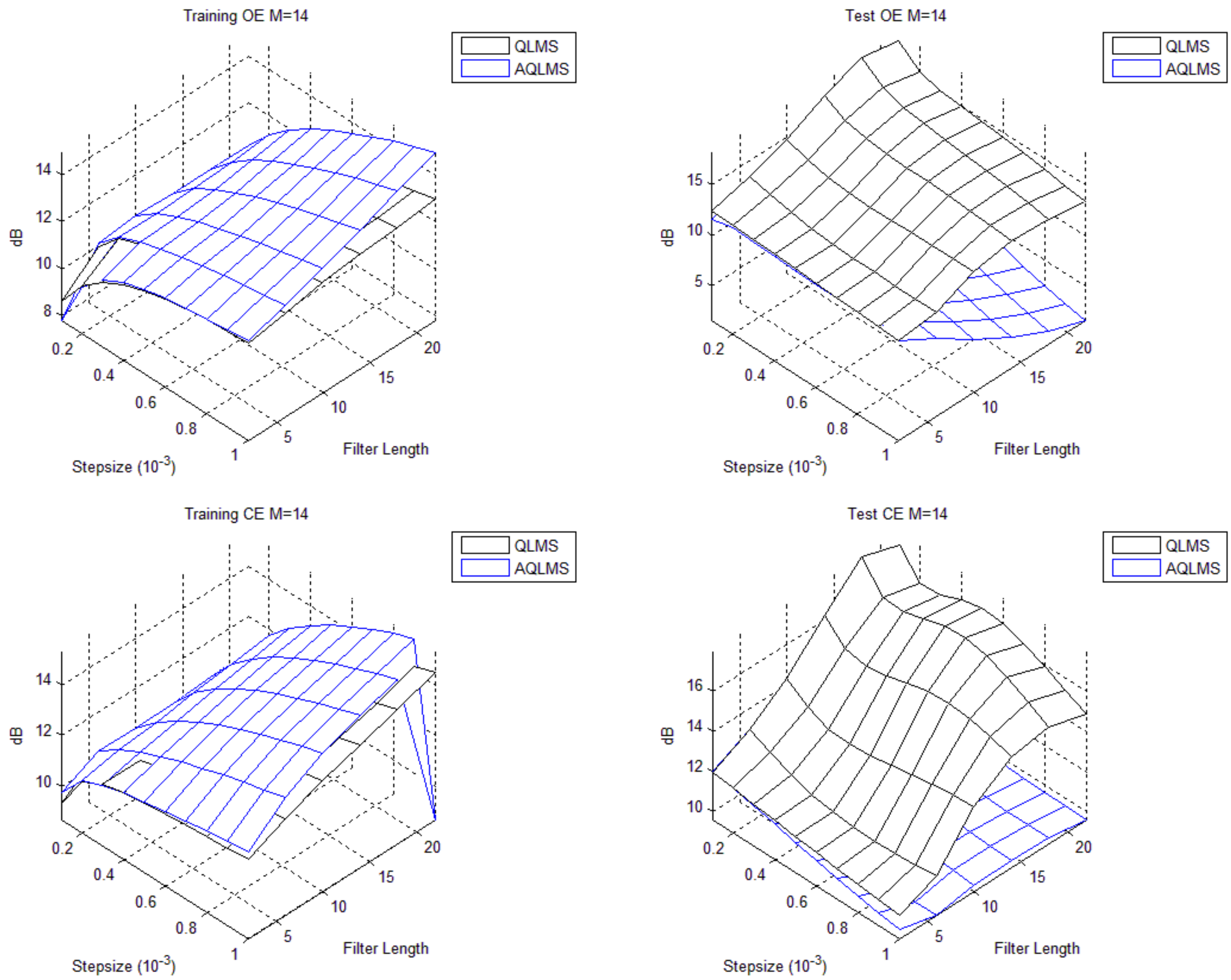


Figure 65 - QLMS and AQLMS performances as function of the stepsize and filter length, for both the training (left) and test (right) phases. The prediction horizon was set to 14 samples (0.14 seconds). In the top are reported the results for the OE case, while in the bottom for the CE case

To better compare the results of the different adaptive filters, we collapsed the results in four unique plots, two for the training OE and CE, and two for the test OE and CE. Starting with the training (Figure 66).

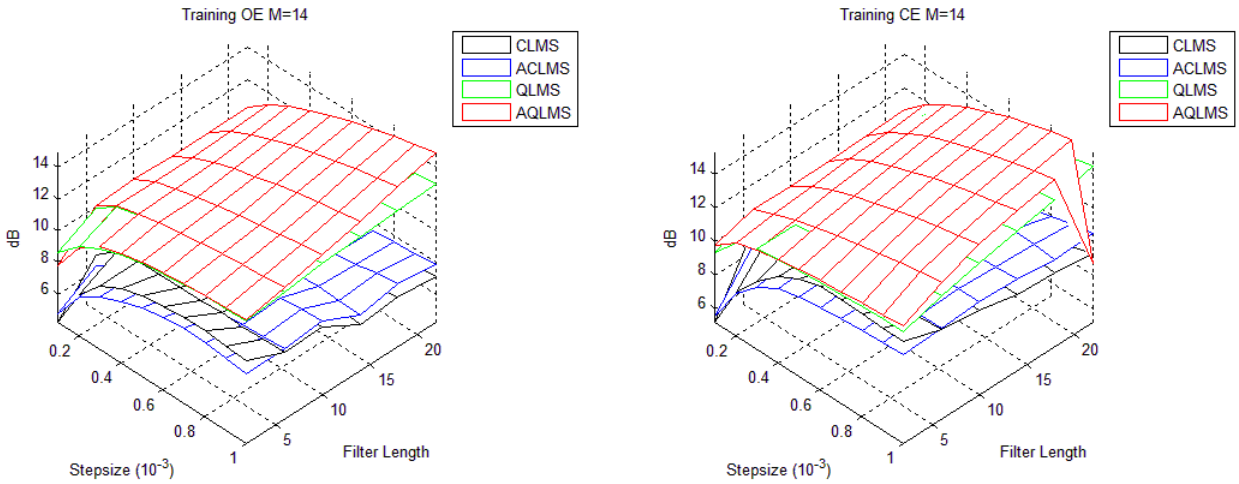


Figure 66 - CLMS, ACLMS, QLMS and AQLMS performances as function of the stepsize and filter length, for the training phase, in the OE (left) and CE (right) case. The prediction horizon was set to 14 samples (0.14 seconds)

The presented results are similar to the ones achieved in (17) for a completely different type of data, a part for the ACLMS, that was not included in their simulations. The order of the performances sees AQLMS, QLMS, ACLMS and CLMS. From these results it is clear that, at least in the training phase, the information shared between OE and CE (quaternion model) and the non-circularity of the signal (augmented statistics) is useful to achieve higher prediction gains. In Figure 67 the results for the test phase were reported.

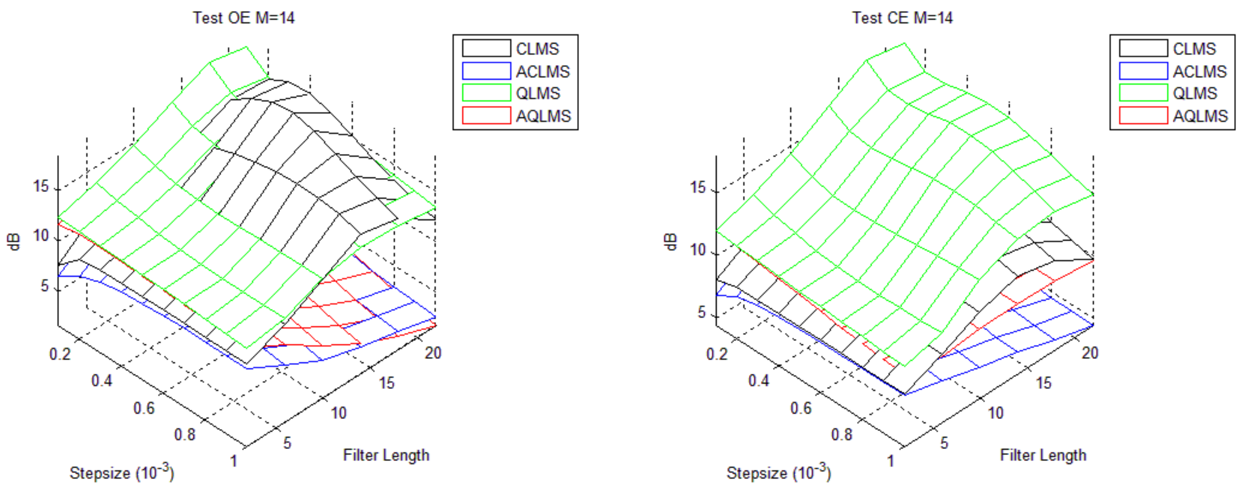


Figure 67 - CLMS, ACLMS, QLMS and AQLMS performances as function of the stepsize and filter length, for both the test phase, in the OE (left) and CE (right) case. The prediction horizon was set to 14 samples (0.14 seconds)

The test phase offers controversial results. The augmented statistics seems to suffer from the overfitting of high order models, obtaining low values of gain with the increase of the filter length. In the OE case the CLMS obtain better performances for a certain range of values of stepsize and filter length, while the QLMS filter performs much better in the CE case. The AQLMS performs better than ACLMS, demonstrating again that in every case, the quaternion models are better than the complex counterparts.

Now we examine the performances as function of the prediction horizon and the filter length. Starting with the analyses of the quaternion filters (Figure 68).

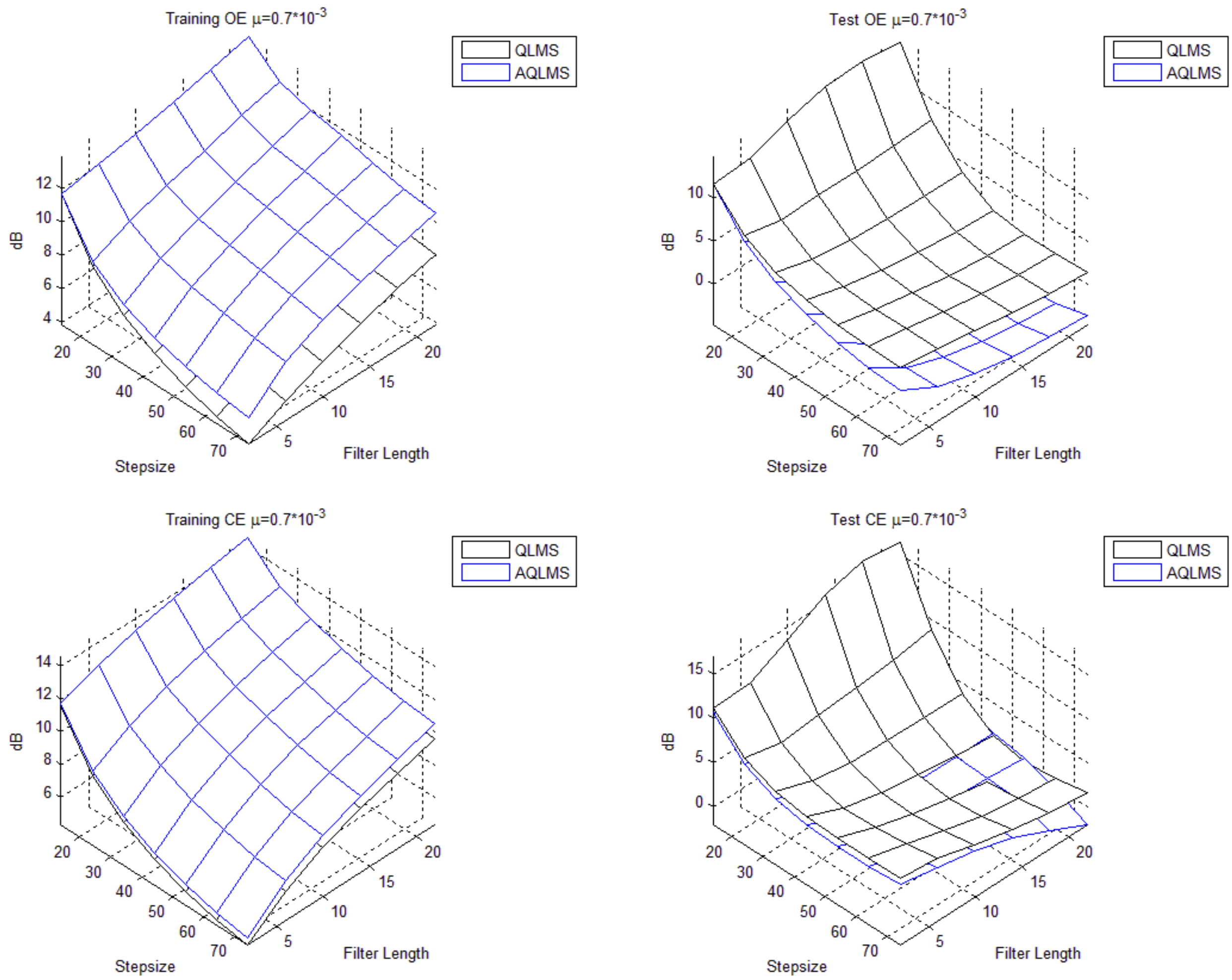


Figure 68 - QLMS and AQLMS performances as function of the prediction horizon and filter length, for both the training (left) and test (right) phases. The stepsize was set to $0.7 * 10^{-3}$. In the top are reported the results for the OE case, while in the bottom for the CE case

It is confirmed also in the quaternion domain that the CE time series achieve a higher prediction gain compared to the OE, in both training and test. The augmented filters perform slightly better in the training phase, and slightly worse in the test phase. We found the same results in the complex domain.

We compared all the results achieved in this type of study (Figure 69).

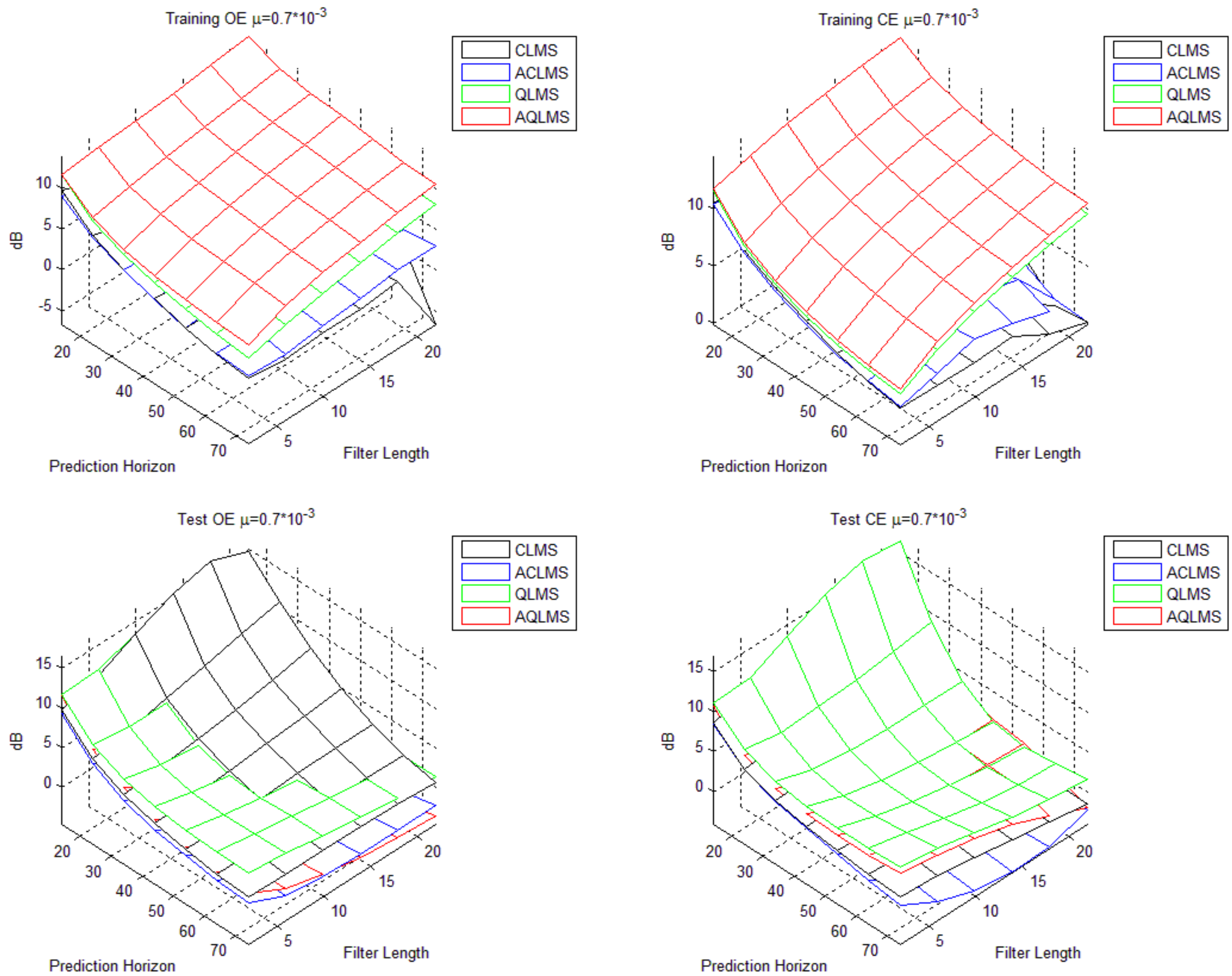


Figure 69 - CLMS, ACLMS, QLMS and AQLMS performances as function of the prediction horizon and filter length, for the training (top) and the test (bottom) phase, in the OE (left) and CE (right) cases. The stepsize was set to $\mathbf{0.7 * 10^{-3}}$.

The achieved results in the training phase follow the same consideration made in the study with the prediction horizon fixed. The AQLMS performs better for both OE and CE (more in the latter case), than there are the QLMS, the ACLMS and the CLMS. The test phase offers different results. Indeed, the OE case follows the precedent study but for the fact that for high prediction horizons the ACLMS performs better than the AQLMS. That would result in the conclusion that for high prediction horizons the complex model is better than the quaternion one. That is not true only in the case of low order filters, indeed in that case the augmented statistics filters perform better than their counterparts. In the CE case, the QLMS performs better, as seen in Figure 67. However, the situation changes for the other filters. Indeed, for high prediction horizons, the AQLMS is the second one to perform better, while previously it was the CLMS. Then the order CLMS, ACLMS is preserved. Therefore for high prediction horizons the quaternion model is better than the complex, and the non-augmented filter are better than the augmented ones.

6. Discussion of the results

The experiments made during the development of this work covered several aspects of the postural sway analysis, reported in the following list

- Evaluation of the complex nature of the CoP time series
- Choice of the best parameter setting for the calculation of the entropy
- Confront between entropy and regularity indexes
- Characterization of the differences related to the visual condition
- Dynamical study of the visual condition
- Dynamical study of the cognitive condition
- Final evaluation of the dimensionality of CoP

Complex nature of CoP

To the best of our knowledge, no one before tried to model postural sway recurring to a complex model. Usually the AP and ML time series are studied separately, or together through the definition of variables like the "resultant distance" (5), that is nothing else than the amplitude of the complex signal obtained studying AP and ML time series together, thus not considering the phase information. As outlined Mandic et al. in (18), the drawback in studying a real bi-variate variable it is exactly to lose the information brought from the phase of the signal, that sometimes could result essential to model properly the system object of the study. The results obtained clearly showed that a CoP model created in the complex domain is statistically different by the real bi-variate model. This study did not allow to qualify which representation is more suited for a certain application. That point was analysed with the study on dimensionality, that will be discuss later.

Choice of the parameters

One of the main problems in nonlinear signal processing is the need to set parameters. Parameter estimation can be solved through optimization methods, but sometimes it is even hard to know which should be the cost function to minimize. In the literature of postural sway, in particular in the computation of the entropy, many are the possible approaches to follow. As outlined during the thesis, the methodology depends on the specific parameter. Three are needed: the time delay, the embedding dimension and the radius. Usually the first one is calculated through the AMI function and the second one through the False Nearest Neighbors (19), the third one relies on different approaches (6), (29). Let's discuss one by one each parameter. We followed a standard approach for the choice of the time delay, using the AMI function. However, we did decide not to choose a value in the interval $[0.01, 0.1]\ s$, as made by many authors ((13), (22), (12)), to explore the effect of larger time delays. Two were the main reasons. The former was the presence of long vertical and horizontal lines in the recurrence plots (not presented in the thesis), that outlined the oversampling of the time series for values within the $[0.01, 0.1]\ s$ interval (even if we used as many authors a sampling rate of 100 Hz). The latter was given by the fractal nature of CoP: expressing fractality in time, our belief is that the fractal nature of CoP imposes to explore time delays in a wider way, not focusing just on short delays. That led us to the choice of 0.2 and 0.5 seconds. The time delays final values and the embedding dimension were assessed through Delay Vector Variance (DVV). We decided not to recur to False Nearest Neighbors, because in literature the values of the embedding dimensions applied to CoP are nearly the same, usually in the interval $[2,6]$ ((22), (29), (6)). Therefore we wanted to explore different ways to solve the parameter settings, keeping the interval $[2,6]$ as a reference. We introduced two parameters linked to DVV that allowed us to better understand the trade-off in the choice of the embedding dimension. Indeed, we found a compromise between the computational load and the values of nonlinearity and stochasticity. The idea under this approach is to choose the embedding dimension to look for a certain property of the time series. Mathematical models are always an approximation of reality, thus it could be wrong to think that a certain parameter could fit for a certain type of dataset: we could have chosen either a small value of the embedding dimension to strengthen the stochasticity and minimize the nonlinearity of the signal, or a large value to strengthen the nonlinearity and minimize the stochasticity. It all depends on the type of study and the researched features of the signal. We recurred to the same approach for the estimate of the radius. Recurring to Recurrence Quantification Analysis we have been able to associate at each radius the relative properties of the time series on which the fuzzy entropy was calculated.

We believe that this approach feature-oriented, even if not optimal, at least it can help to better understand the signal object of study, and its properties.

Confront entropy vs regularity indexes

It is quite tautological that the family of the Approximate Entropy, like Sample Entropy (SE) and Fuzzy Entropy are indexes of complexity (31). In that article, it is explained that complexity is a concept that has multiple descriptions. The complexity of a signal can refer to the unpredictability of a signal, and it can also refer to the difficulties one has in describing or understanding a signal. For example, irregular signals are more complex than regular ones because they are more unpredictable; and regular signals varying quickly appear to be more complex than those varying slowly because quick-varying ones present more variations in a given period of time. This description of complexity means that random numbers are more complex than periodical sinusoidal signals, and that periodical signals with higher frequencies are more complex than those with lower frequencies. With the discover of deterministic chaos, the word "chaotic" stopped to be used to express that a certain phenomenon appeared stochastic, to be replaced exactly by the term "complex". Therefore, complexity and regularity express the opposite concept. $\%Recurrence$ of the Recurrence Quantification Analysis (32) is an index of regularity.

It can happen in literature to find studies where there is a direct comparison between two different works recurring to different methodologies. In (14), for instance, the sample entropy of the CoP of ballet dancers was compared with the ones of the controls (i.e. non-dancers), and the results were confronted with the ones achieved in (13) through recurrence analysis in a similar study (in that case, the controls where track athletes). The authors argued a coherence between the achieved results, because in both the studies it was proven that ballet dancers have respectively higher Sample Entropy (complexity) and lower $\%Recurrence$ (regularity) compared with the controls. However they neglected to cite that the results were discordant respect to the visual condition. Indeed, in both the studies SE and $\%Recurrence$ of ballet dancers were found higher in the OE case. Those considerations proved that there cannot be a direct comparison between two indexes, even if they describe the same features of a system.

In this work of thesis Recurrence Quantification Analysis and Fuzzy entropy (FE) were both used to characterize the first dataset. It was proved that for a certain range of the radius parameter there is an inverse proportionality between $\%Recurrence$ and FE (we compared just AP time series), even if we used different metrics than the ones used in (13) and (14). Thus, we believe that it should be given more attention to the fact that algorithms that calculate a same feature of the signal (entropy, regularity, nonlinearity, etc..) should be not compared without applying them to the same dataset. Indeed, the algorithms rely on mathematical models that could be not so accurate to catch the real essence of the abstract feature of the signal, object of the study.

Considerations on OE vs CE

Recurring to the elaborations with the Fuzzy Entropy we have been able to express our point of view in the evaluation of the entropy of CoP. Precedent studies showed how the estimation of the complexity can lead to opposite results. In (6) the entropy values were found equal in both OE and CE conditions, while in (10) and (29) the entropy was higher in the OE case. Other examples can be taken drawing on the literature of recurrence analysis applied to the postural sway, even if we already discussed that this kind of comparison should be taken as absolutely qualitatively and at high risk of fault. In the RQA literature, there are studies like (12) where $\%Recurrence$ is equal in both the visual conditions, or again in (13) it was shown that track athletes have higher $\%Recurrence$ in the CE case. Interestingly, in the latter article it was demonstrated that using a foam platform instead of a standard one, the $\%Recurrence$ rised for both athlets and ballet dancers in all the visual conditions. What arises from those examples is the dependence of complexity/regularity and other parameters to the physical characteristics of the subjects. Not only. Also the particular conditions of the acquisition, for instance using a foam platform instead of a rigid one, it modifies the value of those parameters.

In our study, we showed the sensitivity of the estimation of complexity/regularity to parameters, presenting the profiles of Fuzzy Entropy and $\%Recurrence$ as functions of the radius for each subject. Those plots demonstrated also that certain subjects have different characteristics compared to the other, no matter which value the radius was. Therefore, some of them had a FE higher in the CE case, while others had similar values of FE in both CE and OE. The situation changed completely when we studied the ML time series. In that case we obtained also examples of FE higher in the OE case than in the CE. The meaning of these results must be contextualized with the considerations made above looking at the literature. It is clear that there is not a specific value for a specific condition. Depending on the physical characteristics of the subject, the training and the conditions of the acquisition, really different values can be obtained. In literature only the conditions of acquisition are considered. The training is considered only when a specific study is being made, like for (13) and (14). Even in our dataset, we did not ask people for their training background. However, we think that for a better understanding of the obtained values, an improvement in the selection process of the sample of study should be made, and specified in the relative article (to allow an easier comparison).

Dynamical study of the visual condition

To the best of our knowledge, no one before tried to study a signal acquired while opening and closing the eyes. Maybe the reason is that it could be somewhat weird: there is actually no need to understand when the subject is in the OE and CE condition, because you can see it. But that is not the real reason to make such a study. The main objectives were the validation of certain parameters as biomarker of the visual condition and the exploration of the changes in the control mechanism under a non stationary regime. Before assessing the sensitivity of the parameters with a traditional windowed study of the time series, we decided to study the sequential correspondence of those parameters. Two are the reasons for that choice. The sequential study is able to show whether a certain parameter can be considered stationary or not, and it can be used to evaluate a minimum number of samples needed for the mathematical convergence of the estimation (there is always an initial transitory due to the small number of samples used in the early stage). We evaluated the standard deviation (variability), the sway-path length (curviness), the fuzzy entropy (complexity) and the Hurst exponent (scaling behavior). Our reference was (10). The results of the standard deviation, the entropy and the scaling exponent of case only-OE and only-CE were coherent in both our study and (10). The sway-path length showed different values, probably due to physical differences between our subject and their sample. When we analised the time series with dynamical changes of the visual condition (test 3), we surprisingly found out that parameters that were able to properly classify the visual condition in the long run (for instance after 60 seconds of records), were not capable to dynamically track it. The only parameters that showed to be more reasonably linked to the visual changes was the sway-path length of the velocity.

Many are the possible interpretation of these results. We would like to differentiate between signals that revealed their stationarity with the sequential study, and signals that did not. In the former case, it would be expected, chosen a proper window length correlated with the transient time of the estimation, to easily track the dynamic changes. Actually that did not happen always. The reason could be due to a differentiation in the control mechanism, that allowed the parameter to converge to a solution in the only-CE or only-OE condition, but did not in the case of OE-CE alternation. In the latter case, it is more reasonable to find something that have problems in tracking the changes. A better approach could be to investigate the time-dependency of those parameters not considering a dynamic scenario, and then contextualize that time-dependency in it.

Dynamical study of the cognitive condition

The final aim of the time-dependent evaluations was to find whether the subjects were altering their postural sway focusing the attention on their body. Again, we used the results of (10) as a reference, even because the only one exploring this type of problem. The authors outlined the entropy as a marker of the attention on the postural control. At the first stages of the work, we tried to find a correlation between their results and a time-dependent study of the Fuzzy Entropy. More precisely we looked for a bistable behavior that could somehow give the impression of a change in the attention/disregard of the body. We found out that entropy was quite far from being stationary (sequential analysis) and used to change presumably too often and within too short intervals (windowed analysis). We argued that a change in the attention could have result in both AP and ML time series, but the correlation coefficient we found, were too small. Because we could not properly understand the meaning of the obtained results, essentially because our dataset was not the most suited to understand the effect of the cognitive component, we created a new one. The parameter which gave the most confident results was the Hurst exponent. In (10) almost all the parameters showed a sensitivity to the cognitive component, but in our case the Hurst exponent was the only one capable to track the cognitive involvement of the subject. In the CE case, it seems that the focusing on the postural sway could be correlated with the presence of the time series in the persistent range. That would be justified by the sensation of oscillation enhancement and memorization that is felt when a person stands still with the eyes closed. With memorization we mean that there is a strong sensation that the movement in a certain direction is amplified, and that sensation is felt for a certain amount of time. Obviously these statements are empirically motivated, but not proven at all. It is a way to interpret the obtained trend.

In this study, the Fuzzy Entropy revealed itself again as the most variable parameter. Those results make us argue that complexity could be a parameter oversensitive to changes. With oversensitive we mean that many changes of the postural system could affect it. Indeed, from the achieved results, Fuzzy Entropy resulted somewhat correlated with both the visual and the cognitive changes.

Quaternion vs Complex based representation of CoP

The result on the complex nature of the CoP obtained at the beginning of the present work motivated the exploration of a dimensionality study. We argued that when the visual condition is missing (i.e. closed-eyes condition), the relevance of the other components that regulate the process of equilibrium is strengthened, as it would happen for a defective system. Indeed biological systems are full of complex regulatory feedback processes that act to maintain their functionality. Therefore, we evaluated the possibility to consider together the information of open-eyes and closed-eyes time series to model more properly the postural control system. To quantify this study, we decided to test whether there was an improvement or not in the prediction gain of an M-sample predictor through adaptive filters. Dependence on filter length, stepsize and prediction horizon was investigated. We recurred to Complex Least Mean Square (CLMS), Quaternion LMS and their Augmented version, ACLMS and AQLMS. The latter two were introduced to take in account the non-circularity of the CoP time series, proved through a specific statistical test. Five are the types of possible considerations:

- Training vs Test
- OE vs CE
- Dependence to the parameters
- Augmented vs non-augmented
- Complex vs Quaternion

Probably because of the short length of the test time series, the training revealed itself as more accurate (i.e. higher prediction gain). Actually we tried to shorten the training time series, and we noticed that the prediction gain is proportional to the length of the time series. We anyway decided to maintain the training with more samples because of the results presented in (16), where it was shown that the ACLMS needs more iterations to converge because of the double number of weights, compared to CLMS. Obviously the reason was to compare the results at the steady state of the different filters, as made in (17) for wind data.

The different visual conditions outlined higher performances of the filters for the closed-eyes case. The CE prediction gains were found higher for both training and test, in all the studies. In particular, the gain differences from OE to CE case for QLMS and AQLMS were small compared to those noticed for CLMS and ACLMS. Those results outlined that a FIR model could be more suitable for the CE case than for the OE one. Moreover, the OE features were not caught by the complex filters, giving poorer results compared to the quaternion filters. Therefore, the CE information helped the adaptive filter to model more properly the OE case, resulting in slight gain differences from OE to CE for quaternion filters, but not for complex ones. This result could be motivated by the fact that in the OE condition the postural sway it is not mainly controlled by the visual information, causing the CE information to improve the prediction.

Apart from the standard considerations in the variation of the parameters, already reported in the Results chapter, the most interesting phenomenon was the improvement for the test dataset of the CLMS gain in the OE case, for both the *Filter Length-Stepsize* and *Filter Length-Prediction Horizon* studies. That improvement arises when the filter length surpasses the 5^{th} order, and it causes CLMS and ACLMS to perform better than QLMS and ACLMS, respectively. The reason could be due again to the problem of the time of convergence of the quaternion filters, which is slower of the complex one. Obviously, that effect rises with the order of the filter, motivating the obtained trends. Another possibility could be that the high order quaternion filters overfit the signal because of the high number of weights, causing a high training, but low test prediction gain. None of these explanations, anyway, explain why it happened only in the OE case.

The non-circularity of the CoP motivated a higher prediction gain of the augmented statistics, as the one reported in the training case. However, during the test the augmented filters performed poorly compared to their counterparts (unless for the case of high order filters). Again, the reason could be one of those explained previously, related to the difference in the number of weights.

Excluding the case presented above, the quaternion filters outperformed their complex counterparts, always. Therefore, as we argued, the hypercomplex model allows the creation of a more accurate model of the postural control system. This result should motivate the passage from a monodimensional bivariate to a multidimensional study of the center of pressure. The introduction of new parameters in the complex/quaternion domain would be needed.

7. Conclusions

During this work of thesis we focused on many aspects related to the study of time series of the Center of Pressure.

We evaluated the setting of the parameters for the reconstruction in the Phase Space of the time series and the calculation of the entropy of the signal. In both the cases, a feature-oriented approach was presented. Thus, we used the Delay Vector Variance and the Recurrence Quantification Analysis to assess the features we were looking for in the time series, to be used in the following elaborations. The decisions were taken considering the, degree of nonlinearity, stochasticity and regularity of the time series.

We studied the differences in the complexity, calculated through Fuzzy Entropy, of the time series under the open-eyes (OE) and closed-eyes (CE) visual conditions. The results obtained motivated the idea that physical differences of the subjects justify different trends in the complexity, explaining that it is wrong to think that complexity should be higher in the OE than in the CE, or vice versa.

A novel dataset was recorded to test particular conditions of stress. In particular we evaluated time series recorded under changes in the visual condition and the cognitive involvement of the subject. We found that the sway-path length is able to classify whether the subject in a certain instant was in the OE or in the CE condition. The Hurst exponent was found as a marker for the cognitive involvement, showing that in the case of distraction from the body sway, the CoP is in the antipersistent range, while in the case of attention on the body sway, it is in the persistent range.

We evaluated the correlation among the anteroposterior (AP) and mediolateral (ML) components of the CoP. We discovered that a complex-based representation of the data is statistically different from a real bivariate. Assuming that the postural control system could have been described more accurately using the information taken from both the OE and CE conditions, we recurred to a quaternion-based representation of the data. Therefore, we joined together the AP and ML time series recorded under the different visual conditions, and we evaluated the prediction gain of a M-steps ahead predictor made through adaptive filters. We evaluated the Complex LMS, the Quaternion LMS, and their augmented versions ACLMS and AQLMS, designed to take in account the non-circularity of the signal, proved through a specific statistical test. We discovered that quaternion filters outperformed their complex counterparts, demonstrating that the information enclosed in the OE records helps the prediction of the CE records, and vice versa. The augmented version of the filters performed better than their counterparts during the training phase, confirming the non-circularity of the signal. However during the test the results were worse, presumably because of the insufficient data length. Indeed, having twice the number of weights, the augmented filters need more iterations to converge to an optimal solution. The CE condition obtained high prediction gains with both complex and quaternion adaptive filters, while the OE condition obtained good results only with the quaternion model. We argued that this result is motivated by the fact that in the OE condition the postural sway it is not mainly controlled by the visual information, causing the CE information to be somewhat correlated with the OE time series, allowing improved predictions.

To conclude, the hypercomplex approach allowed the creation of a more accurate model of the postural control system. This result should motivate the passage from a monodimensional bivariate to a multidimensional study of the center of pressure. The introduction of new parameters in the complex/quaternion domain would be needed.

8. Algorithms

The present chapter explains the different algorithms used during this work, reporting the main concepts and the references for further details.

1. Adaptive Filters

The idea is that the filter adapts its coefficients to obtain certain features. There are mainly four uses of adaptive filtering (36):

- Prediction
- System identification (modeling)
- Equalization (deconvolution, inverse filtering, inverse modeling)
- Interference cancellation

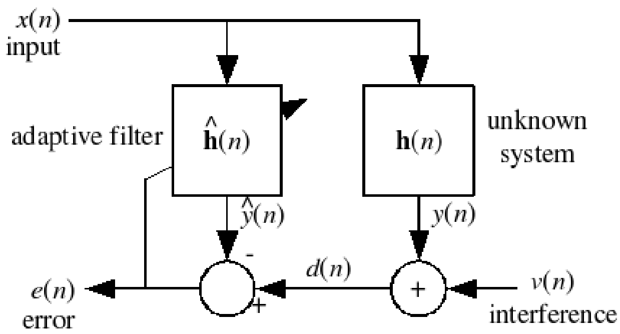


Figure 70 - Example diagram of an adaptive filtering

Usually, the information that updates the coefficients of the filter is the error between a desired value and the actual value. Taking as an example the system identification, reported in Figure 70, we have that the error information is given by:

$$e(n) = d(n) - \hat{h}(n)x(n)$$

where $e(n)$ is the instantaneous output error, $d(n)$ is the desired value (the sum of the input processed by the unknown system and a term of interference) and $x(n)$ is the input data vector of length L.

The update rule used to change the filters weights depends on the strategy used. There are mainly two categories:

- Stochastic gradient descent
- Recursive least-square estimation

In the first case the statistics of the input and between the input and the desired signal are used. In the second case, the actual values of the signals are used.

The Least Mean Square (LMS) belongs to the stochastic gradient descent methods.

1.1 LMS

The LMS update rule can be expressed as

$$\hat{h}(k+1) = \hat{h}(k) + \mu\, e(k)\, x(k)$$

where $w(k)$ is the adaptive weight vector and μ is the step size, the parameter that describes the trade-off between accuracy of the value and response time of the filter and k is the time instant.

The formula arises from the partial derivative of the following cost function

$$C(k) = E\{|e(k)|^2\}, \text{ with } E\{\,\} \text{ denoting the expected value}$$

$$\nabla C(k) = \nabla E\{e(k)e^*(k)\} = 2E\{\nabla\big(e(k)\big)e^*(k)\} = -2E\{x(k)e^*(k)\}$$

and its combination with the descent gradient method

$$\hat{h}(k+1) = \hat{h}(k) - \frac{\mu}{2}\nabla C(k) = \hat{h}(k) + \mu\, E\{e^*(k)\, x(k)\} = \hat{h}(k) + \mu\, e(k)\, x(k)$$

In general, the step size can be selected within the following ranges:

Effect	Range
Stable (convergent)	$0 < \mu < \frac{1}{\lambda}$
Overdamped	$0 < \mu < \frac{1}{2\lambda}$
Critically damped	$\mu = \frac{1}{2\lambda}$
Underdamped	$\frac{1}{2\lambda} < \mu < \frac{1}{\lambda}$
Unstable (not convergent)	$\mu \geq \frac{1}{\lambda}$ and $\mu \leq 0$

Table 16 - Effect of μ on the convergence of the single-weight gradient search process

λ is the trace of the input correlation matrix, defined as $\mathbf{R} = \mathrm{E}[x(k)\, x(k)^T]$.

The learning-curve time constant, that is the number of iterations that the filter needs to follow the input, is given by

$$T_{mse} \approx \frac{1}{4\mu\lambda_n}, \text{ with } n = 0, 1, \dots, L$$

1.2 CLMS

LMS was extended to the complex domain (37) to study bi-variate signals. Complex–valued signals can be either complex by design (communications) or by convenience of representation (radar, sonar). In fact representations in the complex domain provide a natural processing platform, for example processing real domain signals in $\mathbb{C}$ allows the inclusion of phase components, resulting in multidimensional solutions with benefits over real domain solutions (18).

The CLMS algorithm is described by

$$e(k) = d(k) - x^T(k)\, w(k)$$

$$w(k+1) = w(k) + \mu\, e(k)\, x^*(k)$$

Anyway, it does not take into account the differences between the statistics in $\mathbb{R}$ and $\mathbb{C}$. Indeed, the CLMS does not use the pseudo-covariance matrix, assuming implicitly that the input complex signal is circular (for details on the concept of circularity, see 9.2). For this reason the CLMS was improved to ACLMS.

1.3 ACLMS

The Augmented CLMS was created to take into account the statistical properties of non-circular signals. For an example of application to wind forecasting and a comparison with the CLMS, see (38). The improvement stands in the definition of a new input, the augmented complex vector:

$$x_a = \begin{bmatrix} x \\ x^* \end{bmatrix}$$

Having x_a the covariance matrix will be

$$C_{x_a x_a} = \begin{bmatrix} C_{xx} & P_{xx} \\ P_{xx}^* & C_{xx}^* \end{bmatrix}$$

where C_{xx} is the covariance matrix $C_{xx} = E\{x\, x^H\}$, while P_{xx} is the pseudo-covariance matrix $P_{xx} = E\{x\, x^T\}$. For the sake of simplicity the time instant k was omitted.

Note: $(\cdot)^T$, $(\cdot)^H$, $(\cdot)^*$ represent respectively the transpose, the Hermitian and the conjugate operator.

Considering the new input, the update rules for the ACLMS are:

$$y(k) = x_a^T(k)\, w_a(k)$$

$$w_a(k+1) = w_a(k) + \mu\, e_a(k)\, x_a^*(k)$$

where

$$e_a(k) = d(k) - x_a^T(k)\, w_a(k)$$

$$w_a(k) = [h^T(k), g^T(k)]^T$$

with

$$h(k+1) = h(k) + \mu\, e(k)\, x^*(k)$$

$$g(k+1) = g(k) + \mu\, e(k)\, x(k)$$

At the end, the system is

$$y(k) = x_a^T(k)\, w_a(k) = h^T(k)x(k) + g^T(k)x^*(k)$$

For the adaptive filtering of three- and four-dimensional processes, the quaternion least mean square (QLMS) algorithm was introduced.

1.4 QLMS

The Quaternion LMS Algorithm was created for the adaptive filtering of hypercomplex processes (17). By processing data directly in the multidimensional domain where they reside, it is possible to exploit the correlation and coupling between each dimension and therefore provide enhanced modeling. Quaternions can be regarded as a noncommutative extension of complex numbers, and comprise at most four variables (for an introduction, see 9.1).

The explanation is alike the ones seen before:

$$C(k) = E\{e(k)e^*(k)\} = e_a^2 + e_b^2 + e_c^2 + e_d^2$$

$$e(k) = d(k) - w^T(k)x(k)$$

where a, b, c, d refer to the four dimensions.

Doing the derivative of the cost function, and adjusting, it results

$$w(k+1) = w(k) + \mu(2e(k)\, x^*(k) - x^*(k)e^*(k)\,)$$

The term $2e(k)\, x^*(k)$ can be found in the complex LMS (see 8.1.2), but $x^*(k)e^*(k)$ is proper of the quaternion domain.

It is worth to notice that QLMS does not converge to CLMS in the case of null *j* and *k* components. This happens just if quaternion data are in the "isomorphic" form

$$q = q_a + Qi_r \quad \text{where} \quad Q = \sqrt{q_b^2 + q_c^2 + q_d^2} \quad \text{and} \quad i_r = \frac{q_b i + q_c j + q_d k}{Q}$$

1.5 AQLMS

As for the ACLMS (see 8.1.3), the Augmented QLMS was created to process noncircular data (17).

Recurring to a quaternion-valued widely linear model, like

$$y(k) = w^T(k)x(k) + g^T(k)x^*(k)$$

it is possible to incorporate both the information contained in the covariance and pseudo-covariance.

The update rules will be:

$$w(k+1) = w(k) + \mu(2e(k)\, x^*(k) - x^*(k)e^*(k)\,)$$

$$g(k+1) = g(k) + \mu(2e(k)\, x(k) - x(k)e^*(k)\,)$$

Thus, the augmented forms of the update and the input:

$$h_a(k) = [w^T(k)g^T(k)]^T$$

$$x_a(k) = [x^T(k)\;\; x^H(k)]^T$$

Finally, the augmented error and the augmented update can be expressed as:

$$e_a(k) = d(k) - h_a^T(k)x_a(k)$$

$$h_a(k+1) = h_a(k) + \mu[2e_a(k)\, x_a^*(k) - x_a^*(k)e_a^*(k)\,]$$

2. *Delay Vector Variance*

The idea behind this method is to characterize the time series predictability and then compare it with a linearised version of the signal (surrogates data).

Given an embedding dimension m and a time delay τ, a time series X can be represented in the Phase Space, as stated by the Takens' theorem. That transformation creates an m-dimensional representation of the time series, where the coordinates of each point represent a delay vector (DV).

The method can be resumed with the following three steps:

- Mean value μ_d, and standard deviation σ_d are computed over all pairwise distances between DVs, $\|x(i) - x(j)\|$.
- A set Ω_k is generated by grouping all the DVs following the relation

  $$\Omega_k = \{x(i) |\; \|x(i) - x(j)\| \leq \tau_d\}$$

  where τ_d is a threshold given in the interval $[\mu_d \pm n\_sig\; * \sigma_d]$. The number of possible sets Ω_k is given by the number of thresholds τ_d used. The parameter n_alpha defines them, by dividing the interval $[\mu_d \pm n\_sig\; * \sigma_d]$ in n_alpha equal intervals.
- The mean target variance σ^{*^2} is computed over all the sets Ω_k, giving a measure of the unpredictability of the series, that is considered valid by the authors if those sets contain at least 30 DVs.

  $$\sigma^{*^2} = \frac{\frac{1}{N}\sum_{k=1}^{N} \sigma_k^2}{\sigma_x^2}$$

Two plots can be created:

- The DVV plot, were the mean target variance is function of τ_d and is calculated for a range of values dependent on the parameter n_alpha and n_sig
- The DVV scatter diagram, were the horizontal axis corresponds to the DVV plot of the original time series, and the vertical to that of the surrogate time series (obtained averaging the DVV plots of a certain number of surrogates)

Interpretation

About the DVV plot. The presence of a strong deterministic component leads to small target variances for small spans. At the extreme right, the DVV plots smoothly converge to unity, since for maximum spans, *all* DVs belong to the same set, and the variance of the targets is equal to the variance of the time series. If this is not the case, the span parameter, n_sig, should be increased.

About the DVV scatter diagram. If the surrogate time series yield DVV plots similar to that of the original time series, the DVV scatter diagram coincides with the bisector line, and the original time series is likely to be linear. The deviation from the bisector line is, thus, an indication of nonlinearity, and can be quantified by the root mean square error (RMSE) between the ${\sigma^*}^2$ of the original time series and the ${\sigma^*}^2$ averaged over the DVV plots of the surrogate time series (note that while computing this average, as well as with computing the RMSE, only the valid measurements are taken into account). In this way, a single test statistic is obtained, and traditional (right-tailed) surrogate testing can be performed (the deviation from the average is computed for the original, and surrogate time series). In the DVV scatter diagrams, the effect of increasing nonlinearity as described above, corresponds to a stronger deviation from the bisector line. T he span on the horizontal axes of the DVV scatter diagrams becomes smaller as the noise increases.

An example of the two graphs is given in Figure 71.

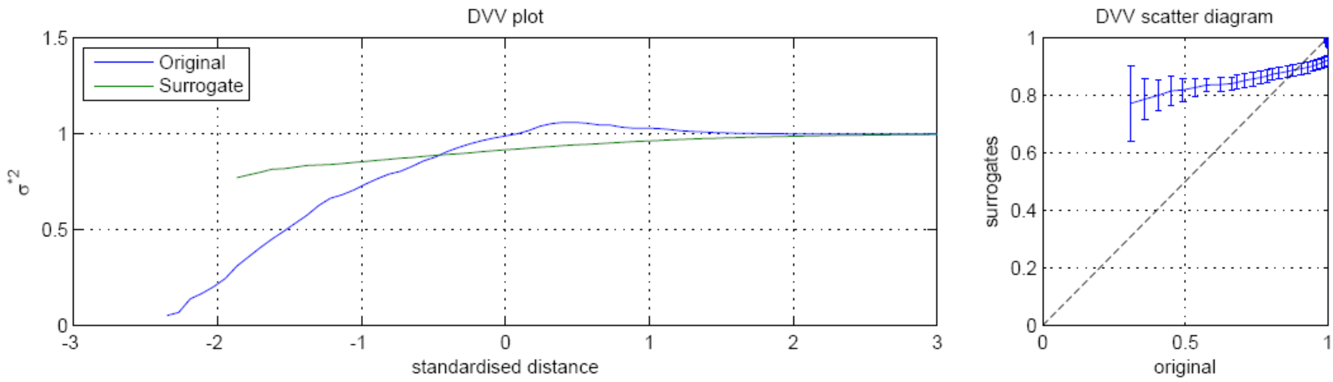


Figure 71 - Example of a DVV plot (left) and a DVV scatter diagram (right), obtained studying a Henon map

2.1 Complex value DVV

The computation of the Euclidean distance between complex-valued DVs is equivalent to considering real and imaginary parts as separate dimensions. Since for bivariate time series, a delay vector is generated by concatenating time delay embedded versions of the two dimensions, the complex-valued and bivariate versions of the DVV method are equivalent.

3. Entropy

3.1 Sample

The Sample Entropy (SE) is a modified form of the Approximate Entropy (39). It is used to characterize the complexity of a time series. Basically, SE is the negative natural logarithm of the conditional probability that a dataset of length N, having repeated itself for m samples within a tolerance ε, will repeat itself for $m+1$ samples, without allowing self matches. Sensitivity to oversampling and the choice of the time delay are a drawback of SE, even if has been proven, according to (29), that SE produces more consistent results than the Approximate Entropy, because it is less sensitive to the length of data. The present algorithm takes into account the improvement of (30), where the time delay was added to the calculation of SE.

After the standard reconstruction in the State Space of the time series y, as specified by the Takens' theorem, the following quantities are calculated

$$\begin{cases} B_i^m(\varepsilon) = \dfrac{1}{N-m*\tau-1}\displaystyle\sum_{j=1,j\neq i}^{N-m*\tau}\Theta(\varepsilon - \|y_i(m,\tau)-y_i(m,\tau)\|_\infty) \\ B^m(\varepsilon) = \dfrac{1}{N-m*\tau}\displaystyle\sum_{i=1}^{N-m*\tau} B_i^m(\varepsilon) \\ B(\varepsilon) = \dfrac{1}{2}(N-m*\tau-1)(N-m*\tau)B^m(\varepsilon) \end{cases}$$

$$\begin{cases} A_i^m(\varepsilon) = \dfrac{1}{N-m*\tau-1}\displaystyle\sum_{j=1,j\neq i}^{N-m*\tau}\Theta(\varepsilon - \|y_i(m+1,\tau)-y_i(m+1,\tau)\|_\infty) \\ A^m(\varepsilon) = \dfrac{1}{N-m*\tau}\displaystyle\sum_{i=1}^{N-m*\tau} A_i^m(\varepsilon) \\ A(\varepsilon) = \dfrac{1}{2}(N-m*\tau-1)(N-m*\tau)A^m(\varepsilon) \end{cases}$$

with:

y the time series
m the embedding dimension
τ the time delay
ε the radius
N the time series' length
Θ is the Heaviside function
$\|\cdot\|_\infty$ the maximum norm

The Sample Entropy is defined as:

$$SampEn(m, \tau, \varepsilon, N) = -log\left(\frac{A(\varepsilon)}{B(\varepsilon)}\right)$$

Choice of the parameters

It is suggested in (29) to use an auto regressive (AR) fitting of the data to chose the m value and then select the ε value so that it minimizes the quantity $Q(m, \varepsilon)$, defined by the following system of equations:

$$\begin{cases} CP(m, \varepsilon) = \dfrac{A^m(\varepsilon)}{B^m(\varepsilon)} \\ \sigma_{CP}^2(m, \varepsilon) = \dfrac{CP(m, \varepsilon)\big(1 - CP(m, \varepsilon)\big)}{B^m(\varepsilon)} + \dfrac{1}{\big(B^m(\varepsilon)\big)^2}[K_A - K_B CP(m, \varepsilon)^2] \\ Q(m, \varepsilon) = max\left(\dfrac{\sigma_{CP}(m, \varepsilon)}{CP(m, \varepsilon)}, \dfrac{\sigma_{CP}(m, \varepsilon)}{-log\big(CP(m, \varepsilon)\big)CP(m, \varepsilon)}\right) \end{cases}$$

where CP is the conditional probability, K_A and K_B are the number of pairs of vectors of dimension $m + 1$ and m that are overlapping, respectively in $A(\varepsilon)$ and $B(\varepsilon)$, and $Q(m, \varepsilon)$ is the maximum relative error.

The choice of the radius ε is based on a $Q(m, \varepsilon)$ that is lower than 0.05. This guarantees that the 95% confidence interval of the entropy estimate is about 10% of its value. The authors outlined that this algorithm is not perfectly applicable to the COP time series, because it overestimates the embedding dimension m. Indeed, estimating entropy of relatively short time series in high-dimensional reconstructed state spaces is difficult and unreliable.

As an alternative method, the authors proposed an empirical approach, developed in two steps:

1. SE is computed for different values of m and ε. The embedding dimension is chosen referring to the SE curves (the study is restricted for the minor values that causes a convergence of the different SE).
2. The median of all the maximum relative errors $Q(m, \varepsilon)$ is computed as a function of ε and the m found at the first step. The final choice of m and ε is given by the minimum error curve.

3.2 Fuzzy

Fuzzy Entropy (FE) is an improved version of the sample entropy, presented in (31). Two are the main changes:

- The time series get detrended, subtracting the mean value from each delay vector
- The Heaviside function is substituted with a fuzzy membership function

Therefore, being y the detrended time series, it is:

$$\begin{cases} B_i^m(\varepsilon) = \dfrac{1}{N - m * \tau - 1} \displaystyle\sum_{j=1, j \neq i}^{N - m * \tau} \mathrm{D_{ij}^{m,\tau}} \\ B^m(\varepsilon) = \dfrac{1}{N - m * \tau} \displaystyle\sum_{i=1}^{N - m * \tau} B_i^m(\varepsilon) \\ B(\varepsilon) = \dfrac{1}{2}(N - m * \tau - 1)(N - m * \tau) B^m(\varepsilon) \end{cases}$$

$$\begin{cases} A_i^m(\varepsilon) = \dfrac{1}{N - m * \tau - 1} \displaystyle\sum_{j=1, j \neq i}^{N - m * \tau} \mathrm{D_{ij}^{m+1,\tau}} \\ A^m(\varepsilon) = \dfrac{1}{N - m * \tau} \displaystyle\sum_{i=1}^{N - m * \tau} A_i^m(\varepsilon) \\ A(\varepsilon) = \dfrac{1}{2}(N - m * \tau - 1)(N - m * \tau) A^m(\varepsilon) \end{cases}$$

$$\mathrm{D_{ij}^{m\tau}} = \mu\left(d_{ij}^{m,\tau}, \varepsilon\right)$$

with:

y the detrended time series
m the embedding dimension
τ the time delay
ε the radius
µ the fuzzy membership function
d the maximum absolute difference of the components of the DVs (as $\|\cdot\|_\infty$ for the SE)
N the time series' length

The fuzzy membership function can be any function with these two properties:

- Being continuous so that the similarity does not change abruptly
- Being convex so that the self-similarity is the maximum

In our elaborations, we used the same suggested in the article:

$$\mu\left(d_{ij}^{m,\tau}, \varepsilon\right) = e^{-\left(\frac{d_{ij}^{m\tau}}{r}\right)^2}$$

We added the use of the time delay in the presented relations.

4. Stationary Process Parameters

For a nice list of parameters to characterize the CoP, refer to (5).

4.1 Sway-path length

It is defined as

$$\mathrm{SP_n} = \frac{\sum|\mathrm{AP[n+1]} - \mathrm{AP[n]}|}{\sigma_{\mathrm{AP}}\,\Delta\mathrm{t}}$$

where Δt is the time involved in the acquisition.

5. Chaotic and Fractal Time Series Parameters

5.1 Average Mutual Information

The Average Mutual Information (AMI) is an important tool to find the proper time delay for the reconstruction of a time-series in the Phase Space, according to Takens' theorem.

It is defined as (19):

$$I(T) = \sum_{s(n),s(n+T)} P\big(s(n), s(n+T)\big) \log_2 \left[\frac{P\big(s(n), s(n+T)\big)}{P\big(s(n)\big)\,P\big(s(n+T)\big)}\right]$$

where T is the time delay, $s(n)$ the time series we are studying, $P\big(s(n), s(n+T)\big)$ the joint probability density and $P\big(s(n)\big)$ and $P\big(s(n+T)\big)$ the individual probability densities of $s(n)$ and its translated version.

The idea is to use $I(T)$ as a kind of nonlinear autocorrelation function to determine when the values of $s(n)$ and $s(n+T)$ are independent enough to be useful as coordinates in a time delay vector, but not so independent to have no connection with each other at all. Usually the first minimum of $I(T)$ is selected as the best time delay.

5.2 Hurst exponent

The approach proposed to calculate the Hurst exponent is taken from (34). The estimation can be resumed using a block diagram Figure 72.

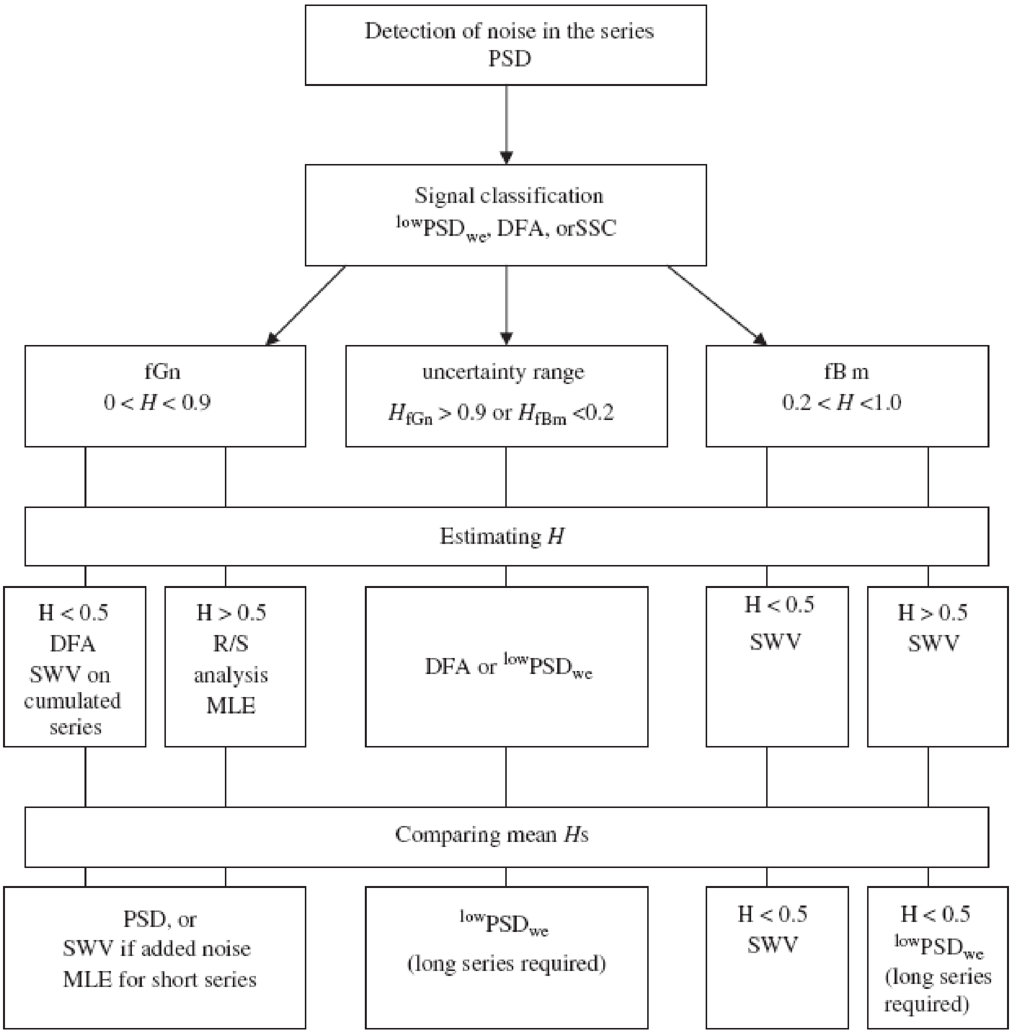


Figure 72 - Block diagram for the estimation of the Hurst exponent

There are mainly three steps:

- The power spectral density (PSD) of the time series is calculated to detect noise in the time series
- A modified version of the PSD, namely lowPSDwe, is used to understand the nature of the time series: if it belongs to the fractional Gaussian noise (fGn) or fractional Brownian motion (fBm) family.
- The proper algorithm is selected depending on the previously obtained Hurst exponent

Below the algorithms used for this work of thesis are resumed. In the article it can be found the characterization of their biases in the estimation, the dependence to the time series length and the sensitivity to noise.

PSD

The power spectral density is calculated using the Wiener-Khinchin theorem, which states that the PSD of a wide-sense stationary process is the Fourier transform of the corresponding autocorrelation function.

$$\begin{cases} PSD(f) = \int_{-\infty}^{+\infty} r_{xx}(\tau)\, e^{-j2\pi f\tau}\, d\tau \\ r_{xx}(\tau) = \mathrm{E}[x(t)\, x^*(t-\tau)] \end{cases}$$

where $x(t)$ is the studied signal and $r_{xx}(\tau)$ its autocorrelation function.

The negative slope of the log-log represents the spectral exponent of the signal, β (it is important to recall that for fractal time series, $PSD(f) \propto \frac{1}{f^{\beta}}$). Then it is needed to check the obtained value:

- If $-1 < \beta < +1$, the signal belongs to the fGn family, therefore $H = \frac{\beta+1}{2}$
- If $+1 < \beta < +3$, the signal belongs to the fBm family, therefore $H = \frac{\beta-1}{2}$

lowPSDwe

It is an improved version of the standard PSD. It is characterized by three pre-processing steps:

- The mean value is subtracted from the time series
- A parabolic window is applied to the data:
  $W(j) = 1 - \left(\frac{2j}{N+1} - 1\right)^2$, for $j = 1,2, \dots, N$
- A bridge detrending is performed by subtracting from the data the line connecting the first and the last point of the series

Then, during the computation of the slope of the log-log plot, the frequencies higher than $\frac{1}{8}$ of the maximal frequency are excluded.

SWV

The procedure for the calculation of the Scaled Windowed Variance follows.

For all the possible interval lengths is repeated:

- The time series $x(n)$ is divided into non-overlapping intervals of length l
- The standard deviation (SD) of each interval is computed, using the mean value of the interval
- The average value $\overline{SD}$ of all the SDs of the intervals of length l is calculated

For a fractal time series, the standard deviation is related to l by a power law:

$$\overline{SD} \propto l^H$$

Therefore, the Hurst exponent is expressed as the slope of the log-log plot of $\overline{SD}$ as a function of l.

There are two main improvements to SWV.

- *Linear detrended* SWV: it is removed the regression line within each considered interval
- *Bridge detrended* SWV: the line connecting the first and the last points of the interval is removed

6. Surrogate Data

The procedures reported below are extensively described in (23). For the CViAAFT, refer to (24).

First, a residual probability α of a false rejection, corresponding to a level of significance $(1-\alpha) \times 100\%$, is selected. Then, for a one-sided test (e.g. looking for *small* prediction errors only), are generated $M = \frac{K}{\alpha} - 1$ surrogate sequences, where *K* is a positive integer. Thus, including the data itself, we have $\frac{K}{\alpha}$ sets. Therefore, the probability that the data by coincidence has one of the *K* the smallest, say, prediction errors is exactly α, as desired. For a two-sided test (e.g. for time asymmetry which can go both ways), are generated $M = \frac{2K}{\alpha} - 1$ surrogates, resulting in a probability α that the data gives *either* one of the *K* smallest *or* largest values. Larger values of *K* give a more sensitive test than $K = 1$. That value is mostly used in order to minimise the computational effort of generating surrogates. Thus, for a minimal significance requirement of 95%, at least 19 or 39 surrogate time series for one-and two-sided tests are needed, respectively.

6.1 iAAFT

Null hypothesis: The time series is generated by a Gaussian linear process, and the only nonlinearity is contained in the measurement function.

The procedure for the creation of the iterative Amplitude Adjusted Fourier Transform (iAAFT) is reported below:

- Computation of the desired number of surrogate time series r, obtained by the permutation of the samples of the original signal, therefore with signal distribution identical.

Then, at every iteration j: (until *error convergence*)

- Computation of the phase spectrum of the surrogate time series: $r \rightarrow \{\phi\}$
- Computation of $s^{(j)}$ as the inverse transform of $\{|S_k| exp(i\phi_k)\}$
- Rank-ordering of $r^{(j)}$ to match $\{c_k\}$

Example of rank ordering:

If $r = [1.1 \quad 3.4 \quad -1.3 \quad 4.5]$, then $rank(r) = [2 \quad 3 \quad 1 \quad 4]$, and represents the indexes of the elements of r if they were ordered in ascending order.

6.2 MViAAFT

Null hypothesis: The time series are generated by a multivariate Gaussian linear process, and the only nonlinearity is contained in the measurement function.

The procedure for the creation of the Multivariate iterative Amplitude Adjusted Fourier Transform (MViAAFT) is:

- Computation of the desired number of surrogate time series r, obtained by the permutation of the samples of the original multivariate signal, therefore with signal distribution identical.

Then, at every iteration j: (until *error convergence*)

- Computation of the phase spectrum of the surrogate time series: $r_1 \rightarrow \{\varphi_1\}$ and $r_2 \rightarrow \{\varphi_2\}$
- Preservation of the cross correlation between the variates, the phases φ are replaced by the phases ϕ, by doing:
  - Calculating the phases of the original time series ρ_1 and ρ_2
  - Calculating the parameter common to both the time series:

$$\alpha_k = \tan^{-1}\left(\frac{\sum_{m=1}^{M} \sin\left(\varphi_{k,m} - \rho_{k,m}\right)}{\sum_{m=1}^{M} \cos\left(\varphi_{k,m} - \rho_{k,m}\right)}\right)$$

$$\phi_{k,m} = \rho_{k,m} + \alpha_k$$

- Computation of $s^{(j)}$ as the inverse transform of $\{|S_k| exp(i\phi_k)\}$
- Rank-ordering of $r_1^{(j)}$ and $r_2^{(j)}$ to match respectively $\{c_{k_1}\}$ and $\{c_{k_2}\}$

6.3 CiAAFT

Null hypothesis: The time series are generated by a multivariate Gaussian linear process, and the only nonlinearity is contained in the measurement function.

The procedure for the creation of the Complex-valued iterative Amplitude Adjusted Fourier Transform (CiAAFT) is reported below. The idea behind the Complex-Valued iAAFT is to extend iAAFT to complex signals.

- Computation of the desired number of surrogate time series r, obtained by the permutation of the samples of the original signal, therefore with signal distribution identical.

Then, at every iteration j: (until *error convergence*)

- Computation of the phase spectrum of the surrogate time series: $r \rightarrow \{\phi\}$
- Computation of $s^{(j)}$ as the inverse transform of $\{|S_k| exp(i\phi_k)\}$
- Rank-ordering of the real and imaginary parts of $r^{(j)}$ to match the real and imaginary parts of $\{c_k\}$
- Rank ordering of the moduli of $r^{(j)}$ to match the modulus distribution of $\{c_k\}$

Note that a rejection of the null hypothesis that the signal is complex-valued and linear, could be due to a deviation from either of the two properties. For this reason a statistical test for the complex nature of the time series was created (see 8.8.1).

7. Recurrence Quantification Analysis

Recurrence quantification analysis (RQA) is a nonlinear and multi-dimensional technique which do not assume data stationarity, which places no restrictions on the statistical distribution of data or on data set length, and which allows the characterization of a variety of features of a given time series.

The basis of RQA is the construction of the time series recurrence plot. A recurrence plot is a matrix created using the following procedure:

- The time series is reconstructed in the Phase Space, according to Takens' Theorem
- A certain metric of interest is selected and then used to calculate the distances between the delay vectors, creating the matrix of the distances
- A certain threshold is selected; the distances within that threshold will be classified as recurrences, the others as non-recurrences

Having the recurrence plot, it is possible to study its features. There is the division in two groups:

Large-scale typologies	*Small-scale textures*
- Homogeneity	- Isolated points
- Drift	- Short line segment
- Periodicity	- Checkerboard

Homogeneity expresses the uniform coverage of the plot of recurrence points. In that case there is the lack of a dynamical structure (an example is noise).

Drift is represented by a trend where the number of recurrence points decreases with the distance from the main diagonal. It expresses nonstationarity in the form of a gradual trend, and in the case of a sudden change in the density of recurrence points, it expresses a rapid change in the level of the dynamics.

Periodicity is represented by diagonal lines parallel to the main diagonal.

Isolated points reflect stochastic behavior.

Short line segments express:

- If diagonal and parallel to the main diagonal, determinism. The length of the segments is inversely proportional to the magnitude of the largest Lyapunov exponent of the signal.
- If diagonal and perpendicular to the main diagonal, there are sequences that are mirrors of other sequences in the time series. Often, this effect is related to the use of a too little embedding dimension.
- If horizontal or vertical:
    - o If the distance is high, the short segments are placed at a certain distance that is more or less the same for the whole dynamic
    - o If the distance is low, the short segments represent a local stationarity of the process

Checkerboard is represented by a mosaic-like appearance of the recurrence plot. It expresses cyclic trajectories of the attractor: Thus, the system passes through near regions of the phase space, and periodically switches between them (as in the Lorentz system, where there is the alternation between the two lobes of the attractor).

As an example the recurrence plots of the AP time series of a CoP, and of a sine wave are reported in Figure 73.

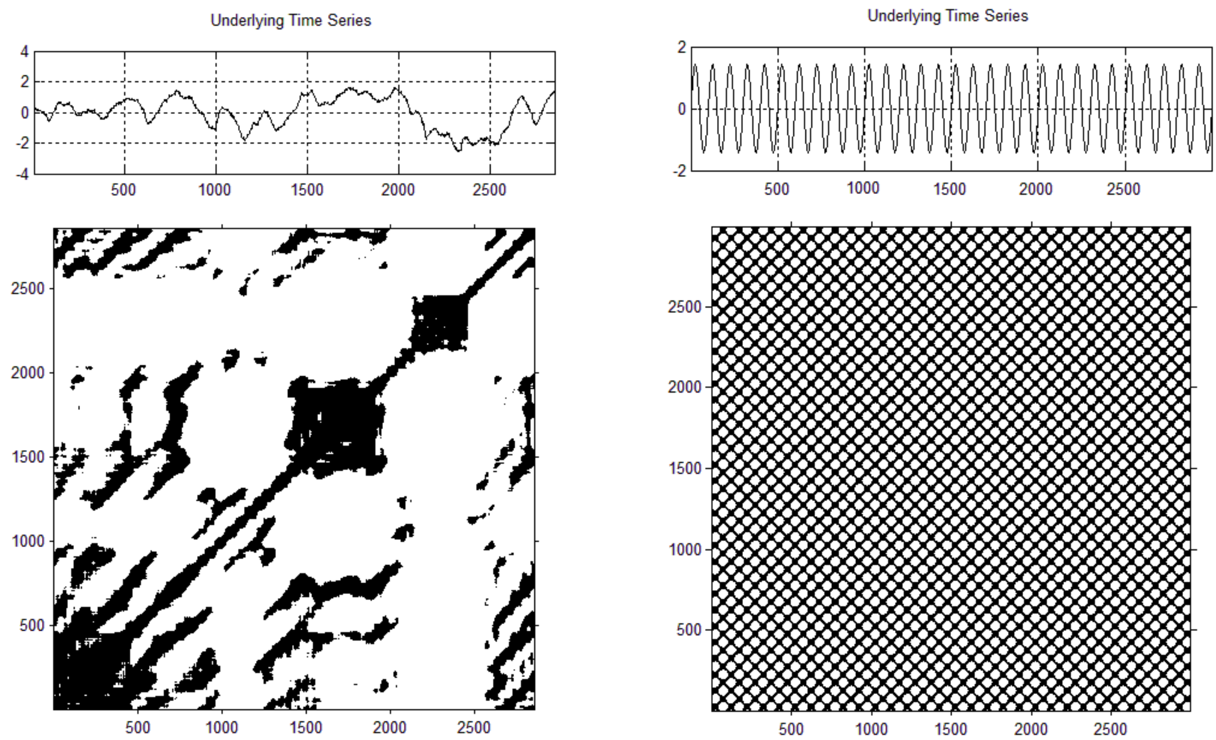


Figure 73 - Example of recurrence plots. AP time series of a CoP (left) and a sine wave (right)

7.1 RQA parameters

$\%Recurrence$ represents the number of recurrences among the total number of possible recurrences. It is calculated as:

$$\%Rec = \frac{1}{N^2}\sum_{i,j}^{N} R_{ij}(\varepsilon)$$

where:

N is the number of delay vectors
$R_{ij}(\varepsilon)$ is the recurrence point i, j of the recurrence matrix. The value is 1 if there is a recurrence, 0 otherwise
ε is the threshold to select the recurrences

$\%Determinism$ represents the number of recurrence points that are part of a diagonal of at least l_{min} elements, divided by the number of recurrences. It is calculated as:

$$\begin{cases} P(\varepsilon, l) = \sum_{i,j}^{N}\left(1 - R_{i-1,j-1}(\varepsilon)\right)\left(1 - R_{i+l,j+l}(\varepsilon)\right)\prod_{k=0}^{l-1} R_{i+k,j+k}(\varepsilon) \\ \%Det = \frac{\sum_{l=l_{min}}^{N} l\,P(\varepsilon, l)}{\sum_{l=1}^{N} R_{ij}(\varepsilon)} \end{cases}$$

where $P(\varepsilon, l)$ is the histogram of the diagonal lines of length l.

$\frac{\%Determinism}{\%Recurrence}$ is a ratio often used to detect transitions of the dynamics. It is calculated as:

$$\frac{\%Det}{\%Rec} = N^2 \frac{\sum_{l=l_{min}}^{N} l\,P(\varepsilon, l)}{\left(\sum_{l=1}^{N} R_{ij}(\varepsilon)\right)^2}$$

7.2 Much more

Recurrence plot and, consequently, RQA can be used also for the elaboration of the joint recurrences between different signals and many others applications. For a complete overview refer to (32).

8. Statistical Test

8.1 Complex-valued nature

The surrogates generated under the null hypotheses of a linear and bivariate time series, and that of a linear and complex-valued time series are compared. The respective surrogates are generated using the bivariate iAAFT (8.6.2) and the CiAAFT (8.6.3). The number N_s of surrogates will depend on the considerations made in 8.6. All time series are characterised using the DVV method (8.2), and a significant difference between the two sets of surrogates is an indication that the original time series is complex-valued.

The proposed methodology is reported below:

- Generate N_s CiAAFT surrogates and the average DVV plot → D_0
- Generate N_s BViAAFT surrogates and corresponding DVV plots → $\{D_b\}$;
- Generate N_s CiAAFT surrogates and corresponding DVV plots → $\{D_c\}$;
- Compare ($D_0 - \{D_b\}$) and ($D_0 - \{D_c\}$).

To perform the final step in a statistical manner, the (cumulative) empirical distributions of root-mean-square distances between $\{D_b\}$ and D_0, and between $\{D_c\}$ and D_0, are compared using a Kolmogorov-Smirnoff (K-S) test. This way, the different types of linearisations (bivariate, $\{D_b\}$, and complex-valued, $\{D_c\}$) are compared to the “reference” linearisation given a complex-valued nature of the time series, D_0. If the two distributions of test statistics are significantly different at a certain level α, the original time series is complex-valued. Therefore, assumptions regarding the possible nonlinearity of the signal are avoided.

8.2 Circularity

The test is presented in (40) and extended to complex non-normal samples (41). The complex-valued measure of circularity based on second-order moments of a complex random variable, called the circularity quotient, is studied (to address the notion of circularity, see 9.2).

Given a zeromean complex random variable $z = x + jy$, we have that z is completely defined by its variance and pseudo-variance, respectively

$$\sigma_z^2 \triangleq E[|z|^2] = \sigma_x^2 + \sigma_y^2$$

$$\tau_z \triangleq E[z^2] = \sigma_x^2 - \sigma_y^2 + j2\sigma_{xy}$$

It is possible to introduce the circularity quotient $\varrho_z \equiv cir(z) \in \mathbb{C}$, defined as:

$$\varrho_z \triangleq \frac{cov(z, z^*)}{\sqrt{var(z)}\,\sqrt{var(z^*)}} = \frac{\tau_z}{\sigma_z^2}$$

ϱ_z can be described as a measure of correlation between z and z^*. The polar representation of the circularity quotient is

$$\varrho_z = r_z\, e^{j\theta}$$

with r_z the *circularity coefficient* of z (i.e. the canonical correlation between z and z^*), and θ the *circularity angle* of z.

Properties and geometrical interpretation of ϱ_z

Let's consider the composite real random vector obtained by the union of the real and imaginary part of z, $v \triangleq (x, y)^T$.

The covariance matrix of v is defined as

$$\Sigma \triangleq E\left[\begin{pmatrix} x \\ y \end{pmatrix}\begin{pmatrix} x & y \end{pmatrix}\right] = \begin{bmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{bmatrix}$$

Its eigenvalue decomposition (EVD) will be determined by the triplet $(\alpha, \lambda_1, \lambda_2)$, respectively the coefficient describing the two eigenvectors $e_1 = \begin{pmatrix} \cos(\alpha) \\ \sin(\alpha) \end{pmatrix}$ and $e_2 = \begin{pmatrix} -\sin(\alpha) \\ \cos(\alpha) \end{pmatrix}$, with $\alpha \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$, and the two eigenvalues.

Those parameters define the ellipse where ϱ_z stands. Indeed, for a positive definite covariance matrix, it is possible to consider the ellipse

$$\mathrm{E}_\Sigma(c^2) \triangleq \{v \in \mathbb{R}^2 : \Delta(v) \leq c^2\}$$

with $\Delta(v) \triangleq v^T \Sigma^{-1} v$, being

In this ellipse, α defines the orientation, c the size, λ_1 and λ_2 define the end points of the major and minor axes. If v has a complex normal distribution, it will be

$$Pr\left(v \in \mathrm{E}_\Sigma\left(\chi^2_{2,p}\right)\right) = p, \text{ where } p \text{ is the } p\text{th-quantile of } \chi^2_2\text{-distribution}$$

The eccentricity of the ellipse will be given by

$$\varepsilon \triangleq \sqrt{\frac{\lambda_1 - \lambda_2}{\lambda_1 + \lambda_2}} \in [0,1]$$

The relationship between the circular quotient and the ellipse parameters is

$$r_z = \varepsilon^2$$

$$\theta = 2\alpha$$

It is possible to link the circularity quotient with the correlation coefficient: (finite nonzero variances are assumed)

$$\rho \equiv cor(x, y) \triangleq \frac{cov(x, y)}{\sqrt{var(x)}\sqrt{var(y)}} \equiv \frac{\sigma_{xy}}{\sigma_x \, \sigma_y} = \frac{\mathrm{Im}[\varrho_z]}{\sqrt{1 - \mathrm{Re}[\varrho_z]^2}}$$

If $\rho \neq 0$, then $sign[\theta] = sign[\rho]$ and $\rho \leq r_z \, sign[\theta] \equiv \varepsilon^2 \, sign[\alpha]$. Hence the circularity quotient lies inside or on the unit circle. The relationships between ρ and ϱ_z are depicted in Figure 74.

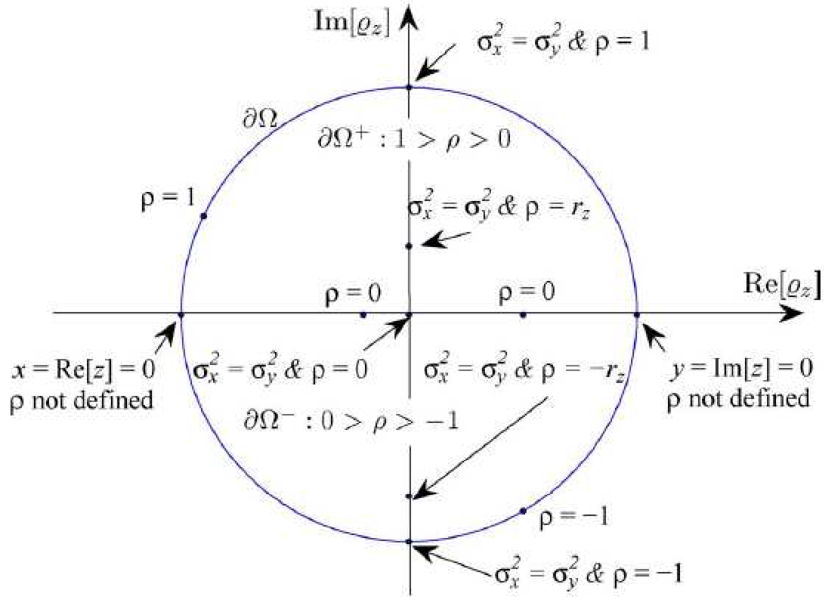


Figure 74 - Pictorial presentation of the properties of ϱ_z

Resuming the results of these analyses, the modulus and phase of the principal square-root of ϱ_z are equal to the eccentricity and angle of orientation of the ellipse defined by the covariance matrix of the real and imaginary part of z. Hence, when the eccentricity approaches the minimum zero (ellipse is a circle), the circularity quotient vanishes; when the eccentricity approaches the maximum one, the circularity quotient lies on the unit complex circle. A connection with the correlation coefficient was established and bounds on ρ given the circularity quotient (and vice versa) are derived.

The test

Assuming that $\{z_i = x_i + jy_i,\ i = 1, \dots, n\}$ is a random normal sample, the Generalized Likelihood Ratio Test (GLRT) of circularity is given by:

$$l_n \triangleq (1 - \hat{r}_z^2)^{\frac{n}{2}}$$

Under the null hypothesis $H_0: \tau_z = 0$ (i.e. the pseudo-variance is null, and the sample is circular), we have that

$$-n \ln l_n \longrightarrow \chi_2^2$$

Non-Gaussian random variables

Assuming that $\{z_i = x_i + jy_i,\ i = 1, \dots, n\}$ is a random sample distributed with a circular Complex Elliptically Symmetric (CES) distribution with finite 4th-order moments. If the null hypothesis is true, there will be:

$$-n \ln l_n \longrightarrow \frac{\gamma}{3} \chi_2^2$$

with γ being the moment of the 4th-order of z. It is therefore possible to define the adjusted GLRT test statistic

$$\ell_n \triangleq -3(n-2)\ln\frac{l_n}{\gamma}$$

9. Appendix

1. Introduction to quaternions

Sir William Rowan Hamilton discovered the quaternions. They are characterized by a scalar part, and a three-imaginary component vector part, defined as

$$q \in \mathbb{H}, \qquad q = \{\mathrm{Re}(q), \mathrm{Im}(q)\} = [q_a, \boldsymbol{q}] = q_a + i\, q_b + j\, q_c + k\, q_d \qquad \{q_a, q_b, q_c, q_d\} \in \mathbb{R}$$

The quaternion vector space $\mathbb{H}$ forms a noncommutative normed division algebra. Their properties are resumed with the following equations:

$$\begin{cases} ij = k \\ jk = i \\ ki = j \\ ijk = i^2 = j^2 = k^2 = -1 \end{cases}$$

A quaternion composed by only its scalar part is said *scalar quaternion*, while if it has only its imaginary part, then it is called *right quaternion* (or vector).

Their applications comprehend three-dimensional rotations, robotics, molecular modeling. Many are the applications also in signal processing, like the spectral estimation with quaternions (Fast complexified quaternion Fourier transform) and the quaternion singular valued decomposition (QSVD).

For references, see (17).

1.1 Quaternions Algebra

The multiplication is given by

$$q_1 q_2 = [q_{a1} q_{a2} - \boldsymbol{q_1} \cdot \boldsymbol{q_2} + q_{a1} \boldsymbol{q_2} + q_{a2} \boldsymbol{q_1} + \boldsymbol{q_1} \times \boldsymbol{q_2}]$$

where $\cdot$ represents the scalar-product, and $\times$ the cross-product.

The conjugate is given by

$$q^* = [q_a, \boldsymbol{q}]^* = [q_a, -\boldsymbol{q}]$$

Therefore, it can be used to extract the scalar and vector parts of a quaternion, according to

$$\mathrm{Re}(q) = \frac{(q+q^*)}{2} \qquad \mathrm{Im}(q) = \frac{(q-q^*)}{2}$$

The *p*-norm is defined as

$$\|q\|_p = (q\, q^*)^{\frac{1}{p}} = (q^* q)^{\frac{1}{p}} = (|q_a|^p + |q_b|^p + |q_c|^p + |q_d|^p)^{\frac{1}{p}}$$

An important property is that

$$(q_1 q_2)^* = q_2^* q_1^*$$

In the quaternion algebra, the square root of -1 is a vector (i.e. pure imaginary) with norm one. The set of all this kind of vectors form the unit sphere.

2. On circularity

A complex random variable $z = x + jy$ is completely defined by its variance and pseudo-variance, respectively

$$\sigma_z^2 \triangleq E\{|z - m_z|^2\} = \sigma_x^2 + \sigma_y^2$$

$$\tau_z \triangleq E\{z^2\} = \sigma_x^2 - \sigma_y^2 + j2\sigma_{xy}$$

Where m_z is the mean value of z, defined as the sum of the mean values of the real and imaginary parts

$$m_z \triangleq E\{z\} = m_x + jm_y$$

For circular random variables $\tau_z = 0$ (i.e. $\sigma_x^2 = \sigma_y^2$ and $\sigma_{xy} = 0$); in that case z would be called *proper*. Therefore non-circularity can be due to two reasons: or x and y have unequal variances and/or x and y are correlated.

Let $\boldsymbol{Z} = \boldsymbol{X} + \boldsymbol{Y}$ be a random vector of $\mathbb{C}^m$. The covariance and pseudo-covariance matrixes are given by

$$\mathcal{C}_z \triangleq E\{\boldsymbol{Z}\,\boldsymbol{Z}^H\}$$

$$\mathcal{P}_z \triangleq E\{\boldsymbol{Z}\,\boldsymbol{Z}^T\}$$

where $(\cdot)^T$, $(\cdot)^H$, $(\cdot)^*$, $E\{\cdot\}$ represent respectively the transpose, the Hermitian, the conjugate and the expectation operator.

The random vector $\boldsymbol{Z}$ is said to be *circular* if

$$\mathcal{P}_z = 0$$

Now, assume that $\boldsymbol{Z}$ is normal, that is $\boldsymbol{X}$ and $\boldsymbol{Y}$ are jointly normal random vectors. Circularity gives many properties to $\boldsymbol{Z}$:

- Rotation invariance: the probability distribution of $\boldsymbol{Z}$ is the same of the probability distribution obtained by $\boldsymbol{Z}e^{j\theta}$, for any $\theta \in \mathbb{R}$
- Given a complex vector $w = u + jv$, the characteristic function is
$$\phi(u, v) = E\{exp[j(u^T\boldsymbol{X} + v^T\boldsymbol{Y})]\} = exp\left[-\frac{1}{4}w^H\mathcal{C}_z w\right]$$
- The odd moments of $\boldsymbol{Z}$ are zero, and the even too if the number of variables involved in the calculation there is not an equal number with of variable with the complex conjugate and variables without it. Indeed, the general form for a moment of order $k$is

$$m_k = E\left\{Z_{i_1}^{\varepsilon_1}, Z_{i_2}^{\varepsilon_2}, \dots, Z_{i_k}^{\varepsilon_k}\right\}$$

with i the indexes of the variables of $\boldsymbol{Z}$, and ε a parameter which values are ± 1.

Relaxing the assumption of normality, the property of rotation invariance leads to the following property:

- Being A and Φ respectively the amplitude and the phase of $\boldsymbol{Z}$, the condition of circularity causes A and Φ to be independent, that Φ is uniformly distributed, and therefore the probability density function of $\boldsymbol{Z}$ is

$$p(A, \Phi) = \frac{1}{2\pi} p(A)$$

with $p(A)$ arbitrary.

For an insight on circularity and other properties see (42).

10. Bibliography

1. *Multisensory fusion and the stochastic structure of postural sway.* **Tim, Kiemel, Kelvin S., Oie and John J., Jeka.** s.l. : Springer, 2002.

2. *A statistical mechanical analysis of postural sway usingnon-Gaussian FARIMA stochastic models.* **Angelo M., Sabatini.** s.l. : IEEE, 2000.

3. **Angelo, Cappello, Aurelio, Cappozzo and Pietro Enrico, di Prampero.** *Bioingegneria della postura e del movimento.* s.l. : Pàtron, 2003. pp. 375-409.

4. *Noise and poise: Enhancement of postural complexity in the elderly with a stochastic-resonance–based therapy.* **M., Costa, et al.** s.l. : EPL, 2007.

5. *Measures of postural steadiness: differences between healthy young and elderly adults.* **Thomas E., Prieto, et al.** s.l. : IEEE, 1996.

6. *Dynamical structure of center-of-pressure trajectories in patients recovering from stroke.* **M., Roerdink, et al.** s.l. : Springer, 2006.

7. *Considerations on the application of the chaos paradigm to describe the postural sway.* **Paolo, Pascolo, Fausto, Barazza and Roberto, Carniel.** s.l. : Elsevier, 2006.

8. *Deterministic center of pressure patterns characterize postural instability in Parkinson's disease.* **Jennifer M., Schmit, et al.** s.l. : Springer, 2006.

9. *Nonstationarities of postural sway.* **Patrick J., Loughlin, Mark S., Redfern and Joseph M., Furman.** s.l. : IEEE, 2003.

10. *Regularity of center-of-pressure trajectories depends on the amount of attention invested in postural control.* **Stella F., Donker, et al.** s.l. : Springer, 2007.

11. *Effects of body lean and visual information on the equilibrium maintenance during stance.* **Marcos, Duarte and Vladimir M., Zatsiorsky.** s.l. : Springer, 2002.

12. *Recurrence quantification analysis of postural fluctuations.* **M.A., Riley, R., Balasubramaniam and M.T., Turvey.** s.l. : Elsevier, 1999.

13. *Dynamic patterns of postural sway in ballet dancers and track athletes.* **Jennifer M., Schmit, Diana I., Regis and Michael A., Riley.** s.l. : Springer, 2005.

14. *Sway regularity reflects attentional involvement in postural control: Effects of expertise, vision and cognition .* **J.F., Stins, et al.** s.l. : Elsevier, 2009.

15. *Complex-valued prediction of wind profile using augmented complex statistics.* **D.P., Mandic, et al.** s.l. : Elsevier, 2008.

16. *Collaborative adaptive filtering in the complex domain.* **B., Jelfs, et al.** s.l. : IEEE, 2008.

17. *The quaternion LMS algorithm for adaptive filtering of hypercomplex processes.* **Clive Cheong, Took and Danilo P., Mandic.** s.l. : IEEE, 2009.

18. *Why a complex valued solution for a real domain problem.* **Danilo, P. Mandic, et al.** s.l. : IEEE, 2007.

19. **Henry D.I., Abarbanel.** *Analysis of observed chaotic data.* s.l. : Springer, 1996.

20. **Edgar E., Peters.** *Fractal market analysis. Applying chaos theory to investment and economics.* s.l. : John Wiley & Sons, Inc., 1994.

21. **Bernard, Picinbono.** *Random signals and systems.* s.l. : Prentice Hall, 1993.

22. *Analysis of postural sway using entropy measures of signal complexity.* **A. M., Sabatini.** s.l. : Springer, 2000.

23. *Surrogate time series.* **Thomas, Schreiber and Andreas, Schmitz.** s.l. : Elsevier, 2000.

24. *A non-parametric test for detecting the complex-valued nature of time series.* **Temujin, Gautama, Danilo P., Mandic and Marc M., Van Hulle.** s.l. : Springer, 2003.

25. *The delay vector variance method for detecting determinism and nonlinearity in time series.* **Temujin, Gautamaa, Danilo P., Mandic and Marc M., Van Hulle.** s.l. : Elsevier, 2003.

26. *Sample entropy analysis of neonatal heart rate variability.* **Douglas E., Lake, et al.** s.l. : American Journal of Physiology, 2002.

27. *Physiological time-series analysis using approximate entropy and sample entropy.* **Joshua S., Richman and J. Randall, Moorman.** s.l. : American Journal of Physiology, 2000.

28. *Comparison of the use of Approximate Entropy and Sample Entropy: Applications to neural respiratory signal.* **Xinnian, Chen, Irene C., Solomon e Ki H., Chon.** s.l. : IEEE, 2005.

29. *On the use of sample entropy to analyze human postural sway data.* **Sofiane, Ramdani, et al.** s.l. : Elsevier, 2009.

30. *Revisiting sample entropy analysis.* **R.B., Govindana, et al.** s.l. : Elsevier, 2006.

31. *Measuring complexity using FuzzyEn, ApEn, and SampEn.* **Weiting, Chena, et al.** s.l. : Elsevier, 2009.

32. *Recurrence plots for the analysis of complex systems.* **Norbert, Marwan, et al.** s.l. : Elsevier, 2007.

33. *Influence of embedding parameters and noise in center of pressure recurrence quantification analysis.* **Christopher J., Hasson, et al.** s.l. : Elsevier, 2008.

34. *Fractal analyses for"short" time series: A re-assessment of classical methods.* **Didier, Delignieres, et al.** s.l. : Elsevier, 2006.

35. *Nonlinear analysis of posturographic data.* **Luigi, Ladislao and Sandro, Fioretti.** s.l. : Springer, 2007.

36. **Bernard, Widrow and Samuel D., Stearns.** *Adaptive signal processing.* s.l. : Prentice Hall, 1985.

37. *The complex LMS algorithm.* **Bernard, Widrow, John, McCool and Michael, Ball.** s.l. : IEEE, 1975.

38. *The augmented complex least mean square algorithm with application to adaptive prediction problems.* **Soroush, Javidi, et al.** Santorini, Greece : IAPR Workshop on Cognitive Information Processing, 2008.

39. *Approximate entropy as a measure of system complexity.* **Steven M., Pincus.** s.l. : Pubmed, 1991.

40. *On the circularity of a complex random variable.* **Esa, Ollila.** s.l. : IEEE, 2008.

41. *Adjusting the generalized likelihood ratio test of circularity robust to non-normality.* **Esa, Ollila e Visa, Koivunen.** s.l. : IEEE, 2009.

42. *On circularity.* **Bernard, Picinbono.** s.l. : IEEE, 1994.