10-4-2016 12:00 AM

Automatic Detection of Eye Blinking Using the Generalized Ising Model

Marwa Elsayh Dawaga, *The University of Western Ontario*

Supervisor: Andrea Soddu, *The University of Western Ontario*
A thesis submitted in partial fulfillment of the requirements for the Master of Science degree in Physics
© Marwa Elsayh Dawaga 2016

Abstract

Electroencephalogram (EEG) is a widely used technique to record electrical brain activity. It is prone to be contaminated by non-neuronal sources that can generate artifacts in the signal due to its sensitivity and its poor signal-to-noise ratio. One of the main challenges in analyzing EEG data is the systematical and effective removal of artifacts from the signal. Although many methods have already been introduced to approach this issue, there is still no robust method for handling all sources of contaminations. For example, eye blinking is a physiological artifact occurring very frequently in spontaneous EEG recordings and therefore, removing these artifacts in a systematic way is a compelling need. The aim of this research is to build an automated pipeline to detect eye blinking artifacts in EEG signals using the generalized Ising model to act as a pattern recognition algorithm. A sample blink pattern is extracted from a single subject whose blink events are validated and marked by an EEG expert. The generalized Ising Model Algorithm works as a fully automated method for identifying all epochs similar to the eye blink pattern. Using the proposed method to discriminate the blinks artifact in continuous EEG data yields optimistic results. From eight healthy subjects, the results show high level of accuracy (90 %).

Keywords: Electroencephalogram, Artifacts, Eye blink, Generalized Ising Model.

Dedications

To my wonderful parents,

For their support and patience…..

To my husband and children

For their unlimited love….

Acknowledgments

The time I spent working on this thesis has been full of incredible moments of learning both on the academic and personal levels. As I am writing those words today, I would like to reflect upon those who were with me through this journey and offered me their help and support.

Firstly, I would like to thank my supervisor Dr. Andrea Soddu for his continuous confidence in me and for providing advice and guidance on every level throughout my Master's program.

I would also like to express my gratitude to Dr. Tushar K. Das (University of Western Ontario), Dr. Robert Camely (University of Colarado) and Dr. Andrea Piarulli (Scuola Superiore Sant'Anna) for their invaluable feedback on my work.

Many thanks to my group members for their encouragement and for many insightful discussions. I am particularly grateful to Pubuditha M. Abeyasinghe for her unlimited support throughout the past two years.

Last but not the least, I would like to thank my husband and my children for their tremendous support and assistance without which the work presented in this thesis would not have been possible.

Contents

Chapter 1: Introduction

List of Figures

List of Tables

List of Appendices

Chapter 1

Introduction

This thesis consists of four chapters. The first chapter provides details about EEG such as the history, recording techniques, applications etc. followed by a brief introduction of EEGLAB and of the Ising Model. The second chapter includes the methodology of implementing our simulations and modifications that allow the simulation to be applied on EEG data. The third chapter presents the results of the analysis. Finally, in the fourth chapter the discussion and the conclusions are presented.

1.1 Introduction

1.1.1 History

Electroencephalogram (EEG) is a widely used technique to record the oscillations of brain electric potential by placing a set of electrodes over the scalp. The first discovery of the electric activity in rabbits' and monkeys' brains, particularly, in cerebral hemispheres was made by the physician Richard Caton in 1875. Subsequently, the neurologist Hans Berger was able to amplify very weak currents recorded from a human scalp, and the name of Electroencephalogram was provided by him. A few years later, the physicians Edgar Adrian and B.H.C. Matthews demonstrated the existence of human brain waves recordings in the Cambridge Physiological Laboratory [1].

1.1.2 Advantages and Limitations

Besides EEG's extensive use among modern medical tools and burgeoning vital role in both clinical diagnosis and cognitive science, it has also brought significant advantages in comparison to other medical techniques.

In general, EEG is relatively inexpensive if we compare it, for instance, to fMRI (functional Magnetic Resonance Imaging). It is also portable, so it is convenient for those who need it. During an fMRI acquisition applied gradients create a very noisy environment inside the machine, which can be uncomfortable for patients and less beneficial for audio studies. On the other hand, EEG is more convenient for audio studies that require silence. Furthermore, EEG is considered as an accessible method for patients who have motor difficulties. Another advantage of the EEG with respect to fMRI is its higher temporal resolution (~1ms), that allows a continuous recording of even short lasting brain activities.

Nonetheless, EEG also has limitations. Spatial resolution of EEG is relatively poor, which makes it difficult to decide which brain portion actually contributed to the EEG signals. Moreover, EEG is hypersensitive to any electric activity and that can result in detection of other undesirable electric activities [2].

1.1.3 EEG and ERP

Event Related Potential (ERP) is the specific usage of EEG to analyze the response of the brain electric potential in events or stimuli. Examples of these events might be images, sounds or both. When the brain receives a stimulus, a series of ERP brain components indicate that different brain systems are actively involved in perceiving and analyzing that stimulus. Because of the high

temporal resolution of EEG, we can understand more about how different information is processed following different stimuli or performing a particular task, and this is considered an extremely strategic aspect for building cognitive neurophysiological tests [3].

1.2 The Source of EEG Signals

The cerebral cortex of the human brain, which is the folded area, has a thickness between 2-5 mm and contains 10^5 neurons per mm^2. A neuron essentially consists of a body cell (soma), an axon and dendrites (Figure 1.1).
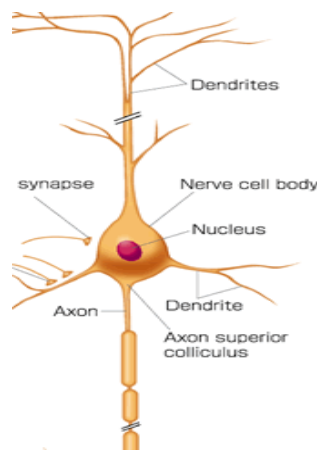
Figure 1.1. The parts of a single pyramidal neuron that connected to three neurons by different types of synapses [4].

Neurons are strongly interconnected with each other by synapses. In the cerebral cortex, neurons named the pyramidal cells, have a unique orientation (perpendicular to the surface of the cerebral cortex) and are responsible for generating the EEG signals [5]. As previously mentioned, neurons interact with other neurons in synapses, and the neurons that receive synaptic inputs are called the target neurons. Excitatory postsynaptic potentials (EPSPs) generated by a synaptic input across cells' membrane cause current sinks (local membrane enclosed by negative charges) in target

neuron and influence it to fire the action potential. On the other hand, inhibitory postsynaptic potentials (IPSPs) are also generated by a synaptic input and cause the generation of current sources. During the synaptic action, micro currents (Na^+, K^+, Cl^-) [6] flow into the target neuron membrane. Due to the relatively long distance between the electrodes and the small currents, the principle of dipole approximation can be applied so that the recorded electric signals are considered as the result of electric dipoles activity. Additionally, only a large number of active pyramidal cells can contribute to recorded EEG signals [5].

1.3 EEG Recordings

In order to record EEG signals, a collection of electrodes needs to be placed over the scalp at specific locations. Systems such as 10/20, 10/10 and 10/5 are international criterions used for positioning the electrodes over the scalp [7]. We provide the following explanations about 10/20 because it is simple and give an idea of the electrodes names based on their location. The 10/20 system is the international criteria used to determine the locations of the electrodes. The numbers "10" and "20" refer to the proportional distances between the electrodes based on the total area of the skull which is determined by the anatomical landmarks "nasion", which is the part of skull between the nose and the "inion" which is the last prominent part of the skull (from front to back), and "ear-lobes" (from right to left) [8].

In Figure 1.3, each channel in this system is identified based on its position, where the letters F, T, C, P, and O refer to the Frontal, Temporal, Central, Parietal, and Occipital locations respectively (See Figure 1.2). Fp denotes the frontal pole while the sub letter z indicates the line-middle electrodes. The even numbers represent the right hemisphere, and the odd ones are set on the left hemisphere (See Appendix B for 2D Plot). The number of electrodes usually used in an EEG are

4

between 16 and 256. Therefore, any additional electrodes can be placed between the labeled 10/20 system electrodes [8]

In addition, referencing is also important, which is the specific choice of one or more electrodes used as the "ground level" for the electric potential. The choice of the reference(s) relies on the type of the required analysis. For an example, if we choose the mastoid as a reference, all the electric activity coming from the electrodes beside this electrode will be subtracted. Thus we have to be aware of whether the information carried from the aforementioned electrodes are relevant for the required analysis or not. Examples for the used references are, the central electrode (Cz), the mastoid and the earlobes (A1 and A2 in Figure 1.3) [1].
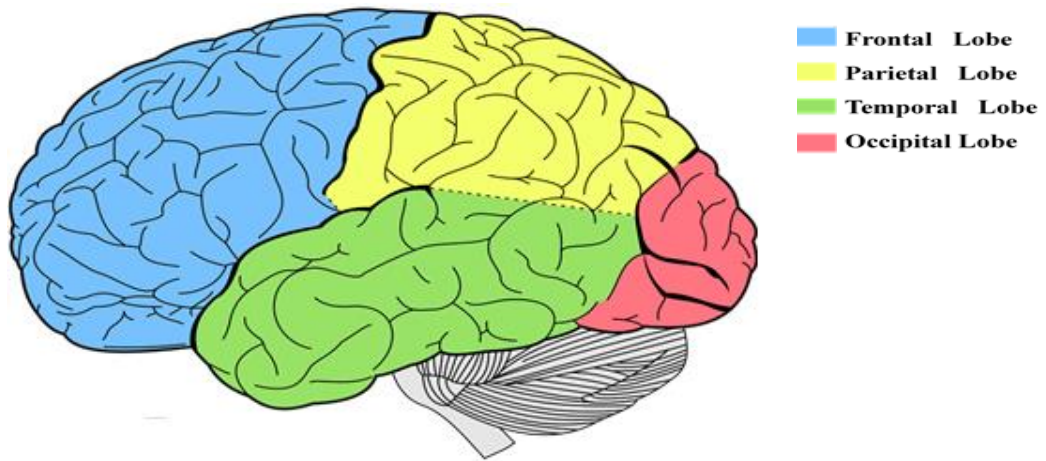
Figure 1.2. The lobes of the brain [9].
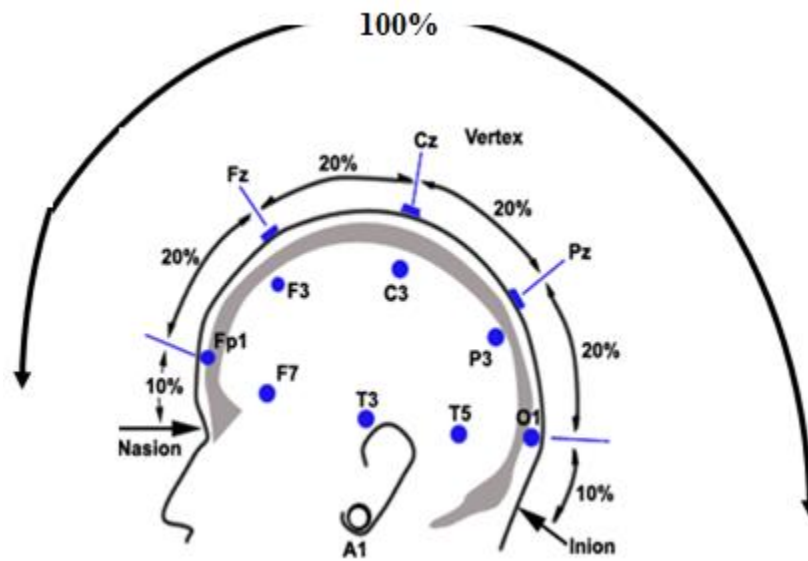
Figure 1.3. The positions of the electrodes based on the international 10/20 system [10].

1.4 EEG Waveforms

The main concern for the cognitive researchers is to link the scalp electric potentials to physiological processes. Therefore, these potentials can be defined based on their temporal and spatial attributes [5]. Specific brain waves can be simply recognized by visual inspection of the EEG signals. These rhythms are also used as signatures of particular diseases or neurological brain states such as dementia and sleep states. [11]. In essence, five main brain rhythms are identified by specific frequency ranges and named based on the Greek alphabet, Delta, Theta, Alpha, Beta, and Gamma.

1.4.1 Delta Waves

In human EEG recordings, delta waves are the waves with the lowest frequencies at < 4 Hz. They appear frequently in recordings of infants, of people with brain injuries and with learning difficulties [12], but most importantly during deep sleep in healthy subjects [13].
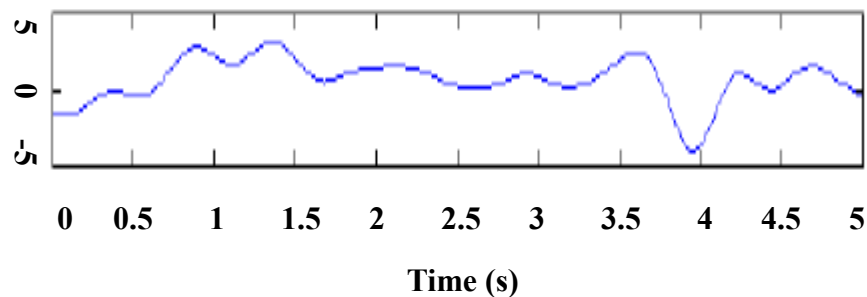
Figure 1.4. Delta brain wave [14].

1.4.2 Theta Waves

Theta waves' frequency lies between 4-8 Hz. They are mainly produced from the frontal region of the brain. In addition, these signals can be associated with sleep disorder and brain tasks that require concentration [15].
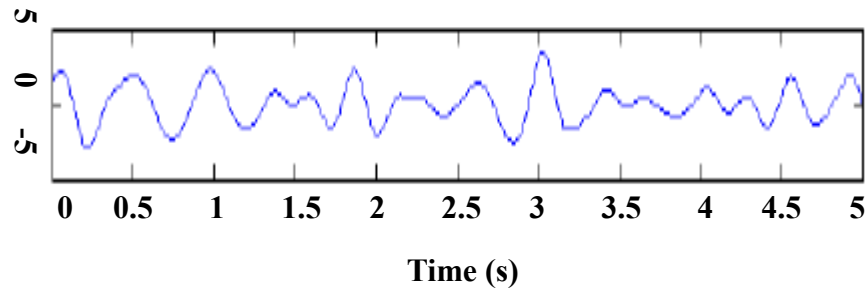
Figure 1.5. Theta waves [14].

1.4.3 Alpha Waves

Alpha was the first discovered and recorded brain oscillation. This waveform can be found in a recording of healthy adults in a state of relaxation and with closed eyes. The frequency ranges between 8 and 12 Hz, and the largest amplitude of Alpha waves can be recorded from the parietal and occipital brain lobes [15].
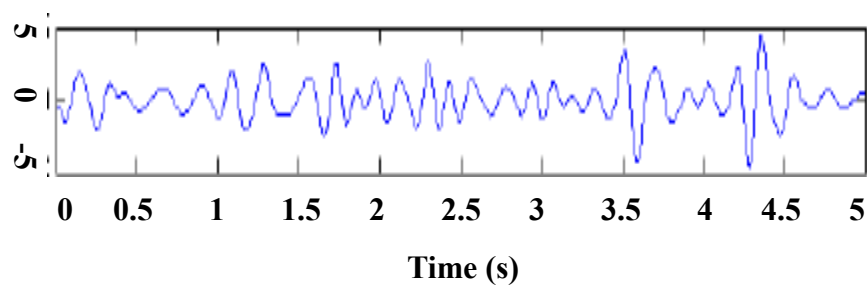
Figure 1.6. Alpha brain wave [14].

1.4.4 Beta Waves

The frequency of beta wave lies between 12 and 30 Hz. It is mostly recorded from the frontal and central areas of the brain and related to concentration and brain tasks. The high level of beta occurs as a result of a participant's stress as well [13].
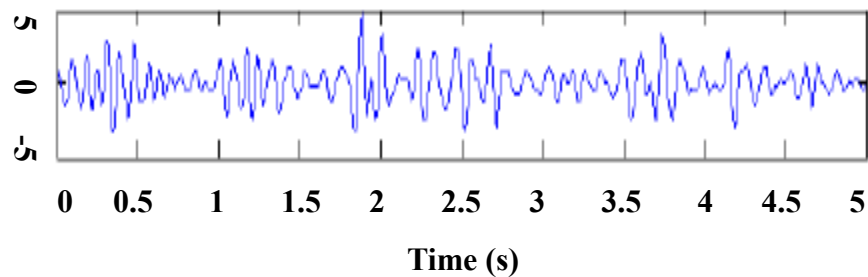
Figure 1.7. Beta brain wave [14].

1.4.5 Gamma Waves

These waves are distinguished from their considered relatively high frequency, 30 Hz and above. Gamma waveform is related to attention. For example, they appear during the sensing of certain stimuli [16].
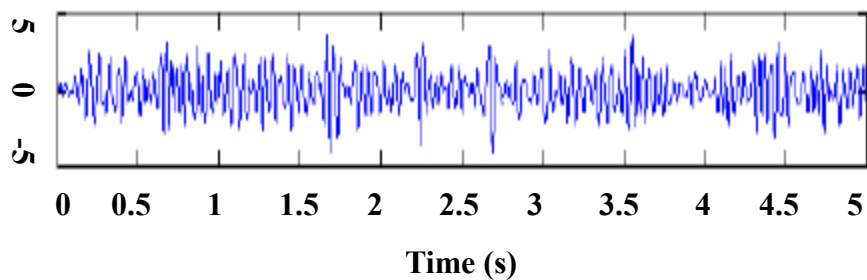
Figure 1.8. Gamma brain wave [14].

1.5   EEG Applications

Some people recognise the EEG cap through its use with epilepsy. Epilepsy is a brain disorder that can be caused by brain injury, illness, or developmental problems though there are also many cases where the causes are still unidentified [17] and consists of an over-excitability of the involved cortex.

 EEG plays an important role in diagnosing and monitoring epilepsy seizures. It is considered the gold standard for diagnosing epilepsy, and has achieved continuous medical development to help patients who experience that disorder [18]. In addition, EEG is used in diagnosing sleep disorders, dementia, brain tumours and brain death. Moreover, EEG is used extensively in cognitive science research, for instance, in those processes that are related to memory, attention, language and emotions because it is a non-invasive technique [1].

1.6 EEG Artifacts

One major obstacle that researchers face in the processing of EEG signals is the artifacts removal as mentioned in Section 1.1.2. EEG artifacts are the electric signals that are not recorded from the brain though EEG was originally designed to record cerebral brain signals. There are essentially two types of these artifacts. The first is known as extra-physiological artifact and the other is physiological artifacts and Figure 1.9 represents EEG signal with different types of artifacts.
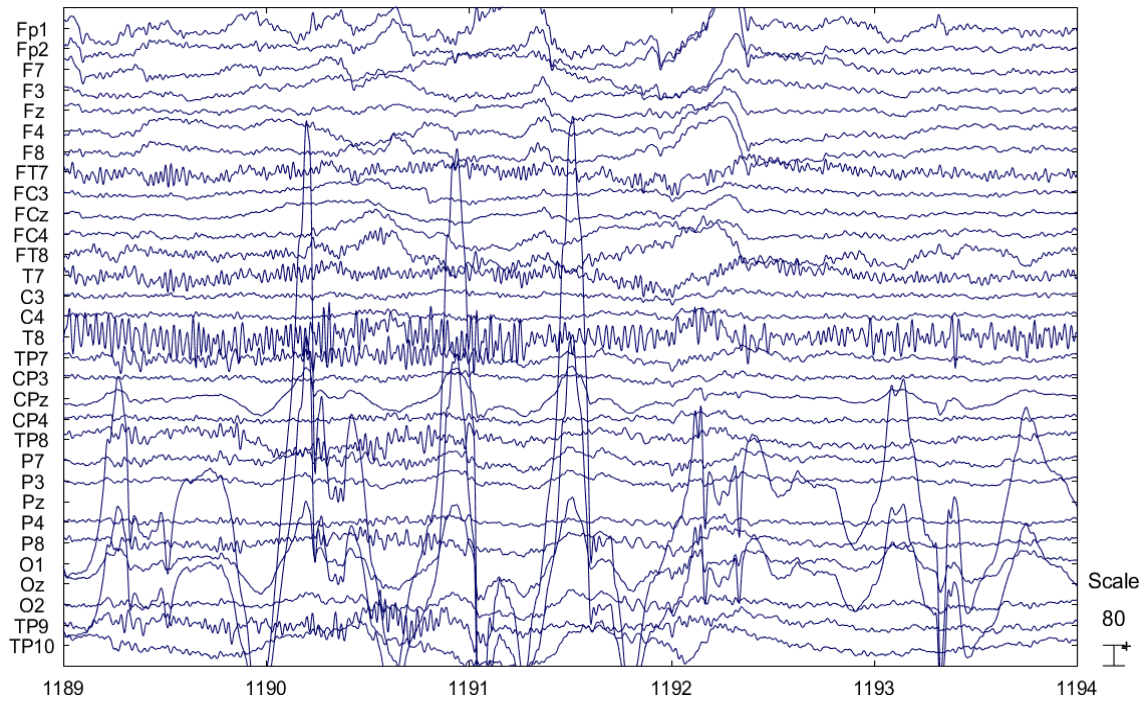
Figure 1.9. Representation of EEG data with different types of artifacts.

Extra-physiological artifacts refer to artifacts of the external electromagnetic signals generated by the power line. Rejecting such signals is not complicated since the discrimination between desired signals and power line signals is straightforward. While the range of cerebral brain signals is 10-100 µVolt, power line signals is 10-1000 mVolts [19]. To remove the power line noise, which is 50 Hz in Europe and 60 Hz in the US, a notch filter filtered on the frequency to remove is applied.

On the other hand, physiological artifacts are the electric signals produced due to the activities of the human body. There are various examples that can be placed under the category of physiological artifacts. However, the following three types of artifacts have been particularly chosen since they occur repetitively in EEG data and several approaches have been introduced in order to eliminate them [20].

1.6.1 Muscle Artifacts

Although the time range for muscle activities is short, it is still important to determine which signals refer to the muscle activity as their frequency range is wide (10 < 500 Hz). The board frequency range implies that the muscle artifacts can be easily correlated with EEG frequency bands. These artifacts also known as electromyogram (EMG) contaminations. To ensure successful recognition and removal for EMG from the EEG recordings, it is better to understand deeply their spectral distribution rather than their frequency [21]. This type of noise can be controlled by the restriction of the subjects' motor activity during the spontaneous EEG recording. However, it is still challenging to eliminate these artifacts in sleep recordings that are mostly contaminated (20-70 % of the all-night recording) [22].

1.6.2 Electrocardiogram (ECG)

These signal are generated by the heart beating potential and often overlap with the EEG signals, especially in subjects with short necks. ECG is a type of signal that can be simply recognized in the EEG recording because it is characterized based on its shape, which appears like a spike, its repetitive occurrence and its lack of correlation with EEG signal. [20]. Due to its relative stationarity, the ECG artefact can be effectively removed using methods such as independent component analysis [23].

1.6.3 Ocular Artifacts

Scalp electrodes are generally distorted by ocular contaminations. However, frontal electrodes are the mainly affected electrodes by these kinds of movements. These contaminations occur as a

result of changes in the electric potential of the eye and the way this electric change is produced depends on the type of eye artifact [24]. In essence, there are two main ocular artifacts, which are the eye movement and the eye blinking.

1.6.3.1 Eye Movement Artifacts

The movement of the eyeball is responsible for these artifacts, and that essentially results in the retino-corneal dipole where the cornea is positively charged with respect to the retina [25]. Based on the direction of the eyeball's movement, we can determine which electrodes are going to be more affected than the others. For instance, if the eyeball is moving vertically, the electrodes such as Fp1 and Fp2 are going to have a greater influence as they are located closely above the eye. On the other hand, the greatest electric potential change will occur to the electrodes F7, in the case of horizontal eyeball movement. [24].

1.6.3.2 Eye Blinking Artifacts

The duration of the human eye blink on average is 0.2 - 0.4 seconds. During the EEG recording, the blink can generate bioelectric signals that have an amplitude 10 times larger than the cerebral signals [25]. Eye blinking creates another type of bioelectric signals that is caused by contacting the eyelids with the cornea; thus the closure of the eye will contribute to attenuation in the EEG signal (Figure 1.10) and the effect can last more than 0.5 seconds (See chapter 2 as we choose the blink duration to be 0.6 seconds) and decreases more quickly the farther the electrodes are. In EEG data, eye blink can contaminate more than 10% of EEG data, which explains the importance of finding an automated way that is capable of separating this type of artifact from EEG data [26].
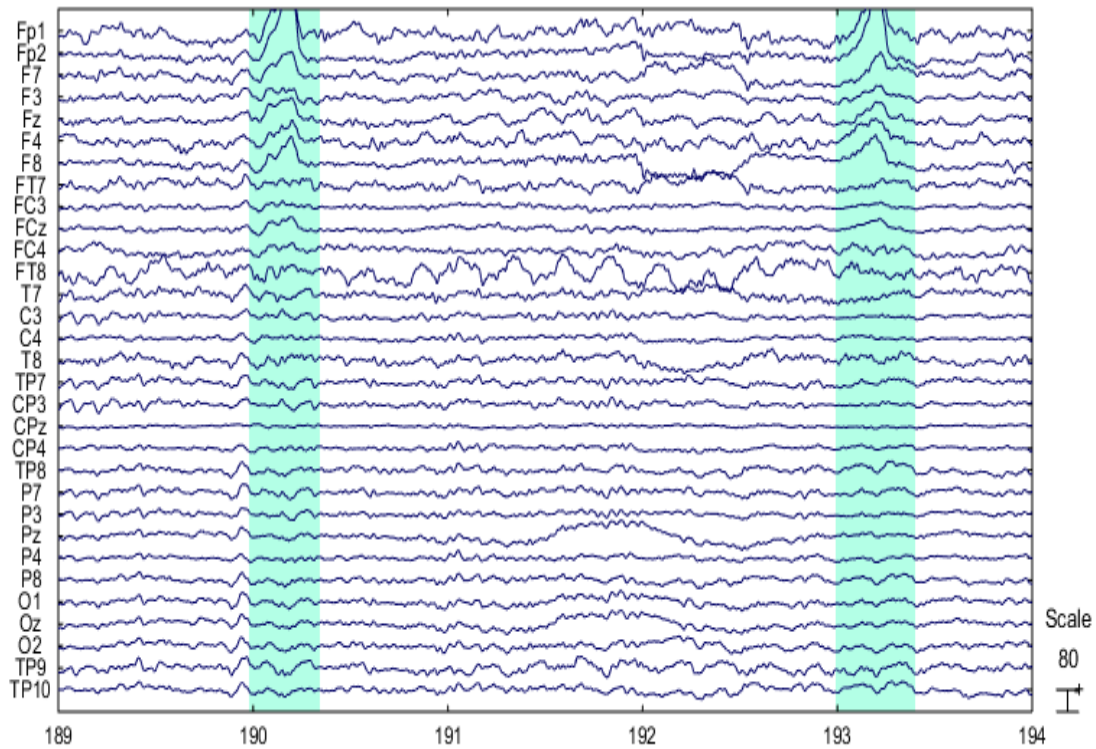
Figure 1.10. Eye blinking artifacts represented by the highlighted areas in 190 and 193 seconds.

1.7   EEGLAB

All EEG analysis was done by using EEGLAB, which is a MATLAB toolbox extensively used in processing EEG, ERP, and MEG (Magnetoencephalogram) data. Specifically, EEGLAB is commonly used to perform the analysis of electrophysiological signals including data visualising, artifact visual rejection, time/frequency analysis (TFA), Independent Component Analysis (ICA) and Graphic User Interference (GUI). ERPLAB is also provided under EEGLAB to study ERPs. Before getting started to use EEGLAB, experience in MATLAB programming is required [27].

1.8   Ising Model

In early 1920s, Wilhelm Lenz studied the phase transition of ferromagnetism in 1-dimintional (chain) spin systems, and that was suggested to be his PhD students' (Ising) proposal [28]. The ferromagnetic materials are known by their spontaneous magnetization due to the alignment of the spins (magnetic moments) in the same direction. Once the temperature of these materials starts increasing, a critical temperature is reached when there is a sudden change of the spins direction. This phenomenon known as phase transition [29].

Ising's first results were not as he expected, and he concluded that there is no phase transition in 1D (1-dimensional) nor in higher dimensions (D > 1) spin systems [30]. However, eventually, scientists proved that there is actually a phase transition in 2D systems.

In 1952, Onsager obtained the exact solution for 2D Ising model. Although it is considered to be a simple model, the use of this model has recently increased in many scientific applications [28] as it is straightforward to solve both 1D and 2D Ising models [31].

The total Energy of state can be described by Equation (1.1) in the absence of an external magnetic field.

$$E = -J \sum_{<i,j>} S_i \ S_j \qquad (1.1)$$

Where E is the total energy, J is the coupling constant, S_i and S_j are the spins in the sites i and j respectively with values +/-1, $< i, j >$ refer to the summation over the nearest neighbors and periodically boundary conditions to reduce size effects are applied.

In this Model, the interactions occur between the spins and its nearest neighbors. Figure 1.11 shows a 2D Ising Model as an example. The number of nearest neighbors for a lattice of size N depends on the systems' dimensionality which can be obtained using 2^d where d is the dimensionality.
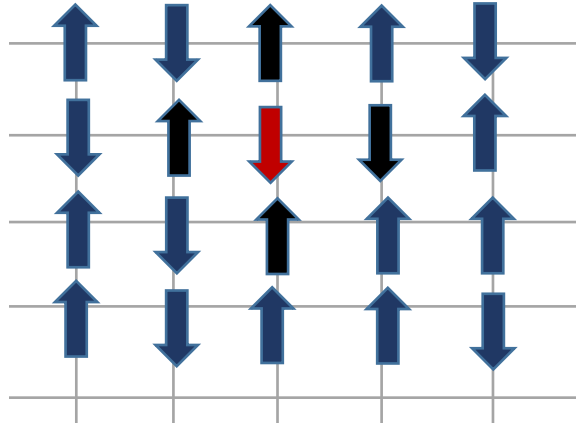
Figure 1.11. Illustration of 2D Ising Model spins. The "red" spin with its four "black" nearest neighbors.

Since we consider the Ising Model as a system in equilibrium with a "heat bath", that implies that we study the system in each equilibrium state (with a fixed temperature T). At the temperature T a configuration "c" is acquired with a probability:

$$p = \frac{1}{Z} \, e^{\frac{-E_c}{K_B T}}$$

$$Z = \sum_c e^{\frac{-E_c}{K_B T}}$$

(1.2)

where Z is the partition function, E_c is the energy and K_B is the Boltzmann constant [32] and the sum in the partition function definition is extended to all possible configurations. Because of the large number of possible configurations (2^N for N lattice sites), it is better to be solved numerically using the Monte Carlo simulations (See Chapter 2).

For the Ising Model, two more representative configurations of spins can be obtained. One, the low energy state or ground state, occurs when the temperature is zero and the spins are aligned in the same direction with an average magnetization of one. The second one can be found at a much higher temperature in which the spins can be directed in one of the two possible direction and the average magnetization is zero. These spin configurations correspond to two distinct phases, which are ferromagnetic and paramagnetic in the case of classical Ising Model. In addition, the state shift can occur related to the temperature of the system: the critical temperature T_c correspond to the barrier between the two phases, and the importance of this temperature is indicated in being a signature of criticality of different usage of Ising model [33].

Chapter 2

Methodology

2.1 Ising Model

The Ising Model is a model that studies the physics of phase transition in ferromagnetism. In 1925, it was introduced by Ernst Ising as his PhD dissertation. The model describes a system that consists of spins (magnetic moments) in a lattice and each spin interacts only with its nearest neighbors with equal couplings across the lattice [34].

Since the model has been used in many scientific applications and can be handled numerically, its actual description for phase transition from ferromagnetic to paramagnetic is called the classical Ising model. Its generalization with non-uniform couplings is usually called the generalized Ising model.

The total energy of the classical Ising model is defined by the following equation:

$$E = -J \sum_{<i,j>} S_i \ S_j - h \sum_i S_i \qquad (2.1)$$

where J is the coupling constant, the notation $< i, j >$ refers to the summation over the nearest neighbors, S_i and S_j are respectively the spins in site i and j, and h is an external magnetic field.

In the absence of the external magnetic field, the energy of the system becomes:

$$E = -J \sum_{<i,j>} S_i \ S_j \qquad (2.2)$$

the spins take either the value +1 for the upward direction or -1 for the downward direction; therefore, from Equation (2.2) we can deduce that as the spins are aligned in the same direction (ordered spins), the energy will have the lowest energy value.

In this research, we are considering the generalized Ising model where the J (coupling constant) is going to be different from the classical coupling, and the choice of J_{ij} is set in order to use the generalized Ising model as a pattern recognition tool.

Let's assume for a 10x10 square lattice, the J_{ij} can be defined as in Equation (2.3) for a well-defined configuration as, for instance, in the letter "A" case (see Figure 2.1). As shown in Figure 2.1 the upward spins are visible, while the downward spins are represented by blank sites.

$$J_{ij} = S_i(A)S_j(A) \qquad\qquad (2.3)$$
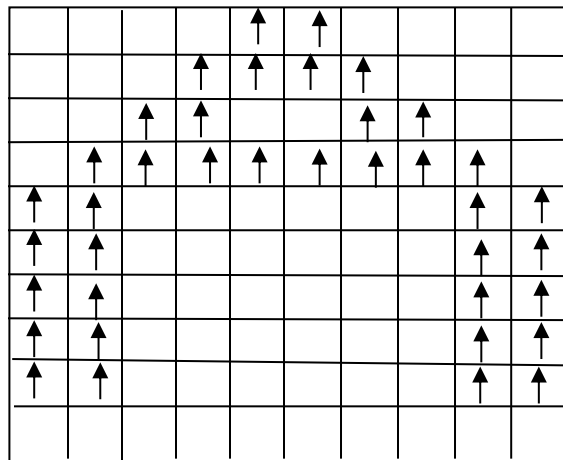
Figure 2.1. Spin configuration of letter "A" represented by upward spins and with the rest of the spins downward.

.

Based on the assumption that $J_{ij} = S_i(A)S_j(A)$, we can calculate the energy as:

$$E = -\sum_{i,j} J_{ij} \, S_i \, S_j \qquad (2.4)$$

which for the particular 'A' spin configuration gives

$$E = -\sum_{i,j} \left(S_i(A)S_j(A) \right) (S_i(A)S_j(A)) \qquad (2.5)$$

Since the S_i and S_j can be either +1 or -1, and both i and j go from 1 to the number of lattice sites (N), then :

$$E = -N^2. \qquad (2.6)$$

This value of energy is the lowest possible energy, and this is important as we can then identify the letter "A" as the lowest energy configuration. We can also have the same energy from a configuration of $\{- S_i(A)\}$, which is going to be the inverse of the latter "A" and that means the downward spins represents the letter "A" and the rest of spins in the configuration are upward.

Let us consider now another initial configuration similar to the letter "A" spatial pattern, for example "À", as in Figure 2.2. Being the energy of the "À" configuration higher than the energy of the letter "A", the pattern "À" will have the tendency to converge to the most energetically favourable configuration with pattern "A". It is a common procedure to compel the evolution of

an initial configuration to a more favourable one by randomly flipping spins in a Monte Carlo simulation (explained in Section 2.2).
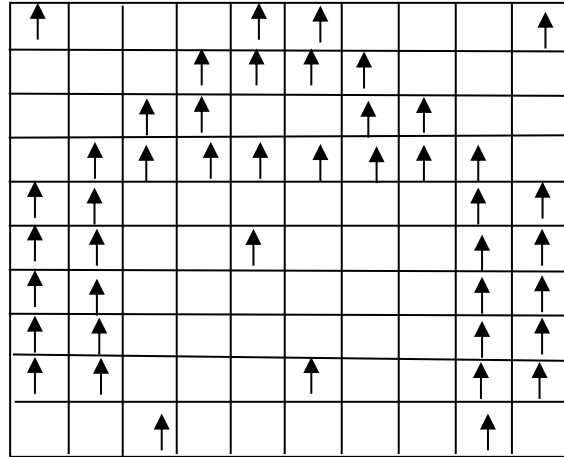
Figure 2.2. The representation of "À", a spin configuration similar to the letter "A", which can be obtained by perturbing the letter "A".

2.2 Monte Carlo Simulation

Monte Carlo methods are computational techniques commonly used in scientific applications in order to obtain a valid approximation of the solution of the problem in a reasonable computational timing. This technique allows for the estimation of the statistical properties of the Ising model by generating the more likely configurations of the system [35].

Considering the configuration in Figure 2.2 as the starting configuration, we can describe the Monte Carlo simulation in the following steps:

1. Choose a spin randomly to flip.

2. Calculate the energy change ΔE which occurred by flipping the selected spin.

3. Check the condition $\Delta E < 0$, and if the condition is satisfied then accept the flip.

4. If $\Delta E > 0$, then accept the flip with probability p=exp $(-\Delta E/KT)$; otherwise, reject the flip and return to the original configuration.

5. Finally, to obtain the equilibrium energy, repeat the previous steps for a number of required trials.

2.3 Blink Detection in EEG Data

2.3.1 Subjects

EEG data acquired from eight healthy subjects, with mean age 46± 15, including three females. (Appendix C).

2.3.2 EEG Data Acquisition

For EEG data acquisition, the system that composed of 256 is used to record EEG data (For the purposes of our research, only 32 channels have been selected and the recordings obtained by using 250 Hz as a sampling rate.

2.3.3 EEG Data Simulation

In the previous section, we discussed the potential convergence of an initial configuration to a chosen pattern.

In this section we introduce the possibility of using the generalized Ising model as a tool to detect blinks in spontaneous EEG data. A two dimensional pattern for the blink is going to be built by using the number of electrodes and time respectively.

As we explained previously, the spontaneous EEG data have been obtained from eight healthy subjects (mean age 46± 15, 3 females) by selecting 32 electrodes from a 256 net where the Cz is used as the reference electrode. From the eight subjects, the first two subjects' continuous recording EEG were used to extract blink events (extraction was made by an EEG expert). The blink events extracted from the first subject were employed to build an average blink pattern, while the blink events extracted from the second subject were used as the initial configurations. Then their convergence to the average blink pattern was tested.

As described in section 1.6.2, because the eye blinking mostly affects the frontal electrodes, we decided to choose only the electrodes that are mostly impacted by the blinking in order to build the blink pattern. Also because the blink attenuates the EEG signals for more than 0.5 seconds, we set to 0.6 seconds the time to represent the blink duration. The average blink was built from 12 out of the 32 channels and from 150 time points, which correspond to a 0.6 seconds' time window with a 250 Hz sampling rate. We considered the first 12 channels, the frontal and prefrontal channels (Fp1, Fp2, F3, F4, Fz, F7, F8, FT7, Fc3, Fcz, Fc4, FT8) as when we used the EEGLAB to visualize the data, these channels were obviously affected by blink activity more than the other channels. The average was performed over the 25 blinks extracted by an expert when looking at the continuous recording of the first subject. Figure 2.3 shows the average blink for the first channel.
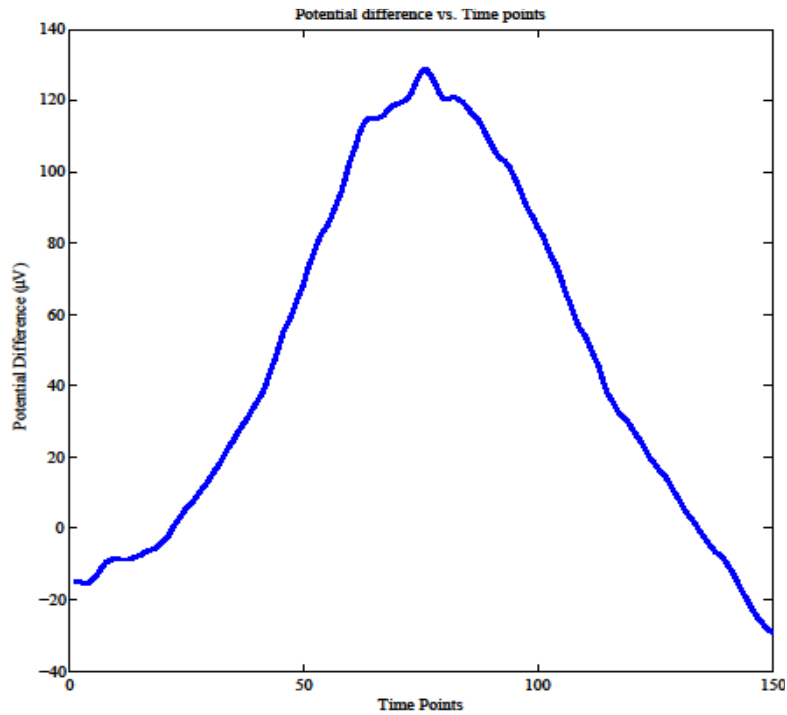
Figure 2.3. Average blink time profile when looking at the first channel.

For the second subject the expert isolated 84 blinks. As mentioned above, the average blink is a 12x150 matrix, and to make it more computationally friendly for the simulation, the 150 time points were down sampled to 30 in order to get the final average blink pattern with a 12x30 matrix. Finally, in order to create a digital pattern, either +1 or -1, which could be used in the generalized Ising model simulations, a convenient baseline of 14 µVolts was empirically introduced.
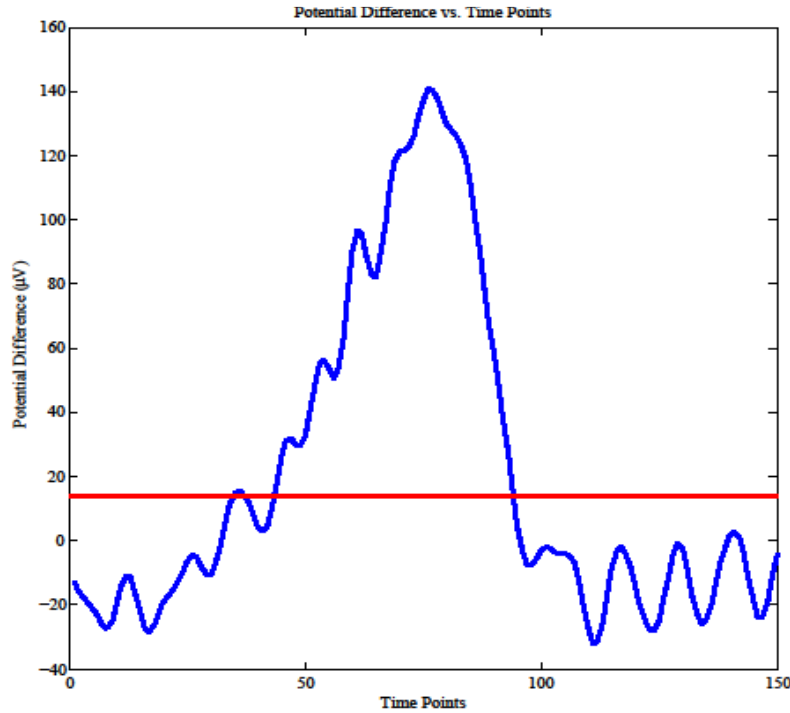
Figure 2.4. A blink from one channel with chosen baseline.

The same baseline for all the 12 channels has been adopted and a value of +1 and -1 have been assigned for a signal respectively greater and lower than the baseline producing a 12x30 blink pattern of ±1.

Using exactly the same procedure as for the letter A, the average blink pattern configuration has been subsequently adopted to build the couplings $J_{ij}(Blink) = S_i(Blink)S_j(Blink)$ with the energy as in Equation (3) for the letter A. When focusing on the 84 recognized blinks from the expert, each of the 84 blink configurations was prepared for simulation after down sampling from 150 time points to 30 and digitizing the signal. Each blink configuration went through the above described Monte Carlo steps, and the final configuration where each blink configuration converged was then compared to the average blink configuration. When convergence to the average blink

happened, full overlap or zero distance between the two configurations, we could claim that the blink was recognized by our detector.

What has been implemented so far has required an expert to identify the blinks, isolate them, and only then we tested if our recognition procedure was successful or not. Of course if we want to build a blink detector we need to be able to move along the full time window of acquired data and have our detector being able to recognize the blinks whenever they happen. In the case of continuous EEG recordings, the challenge is that we might not be able to select the entire blink if the simulation is taking a 150 time points as step; thus, we choose to shift by 10 time points instead of 150. Then using a slicing pattern, we adopted the same procedure implemented of down sampling, digitizing, and comparing with the averaged blink pattern. Once the blink is recognized in the EEG dataset, a flat line of zeros replaced the entire 12x150 matrix. Figure 2.5 illustrates the technique of shifting with 10 time points by considering one channel.
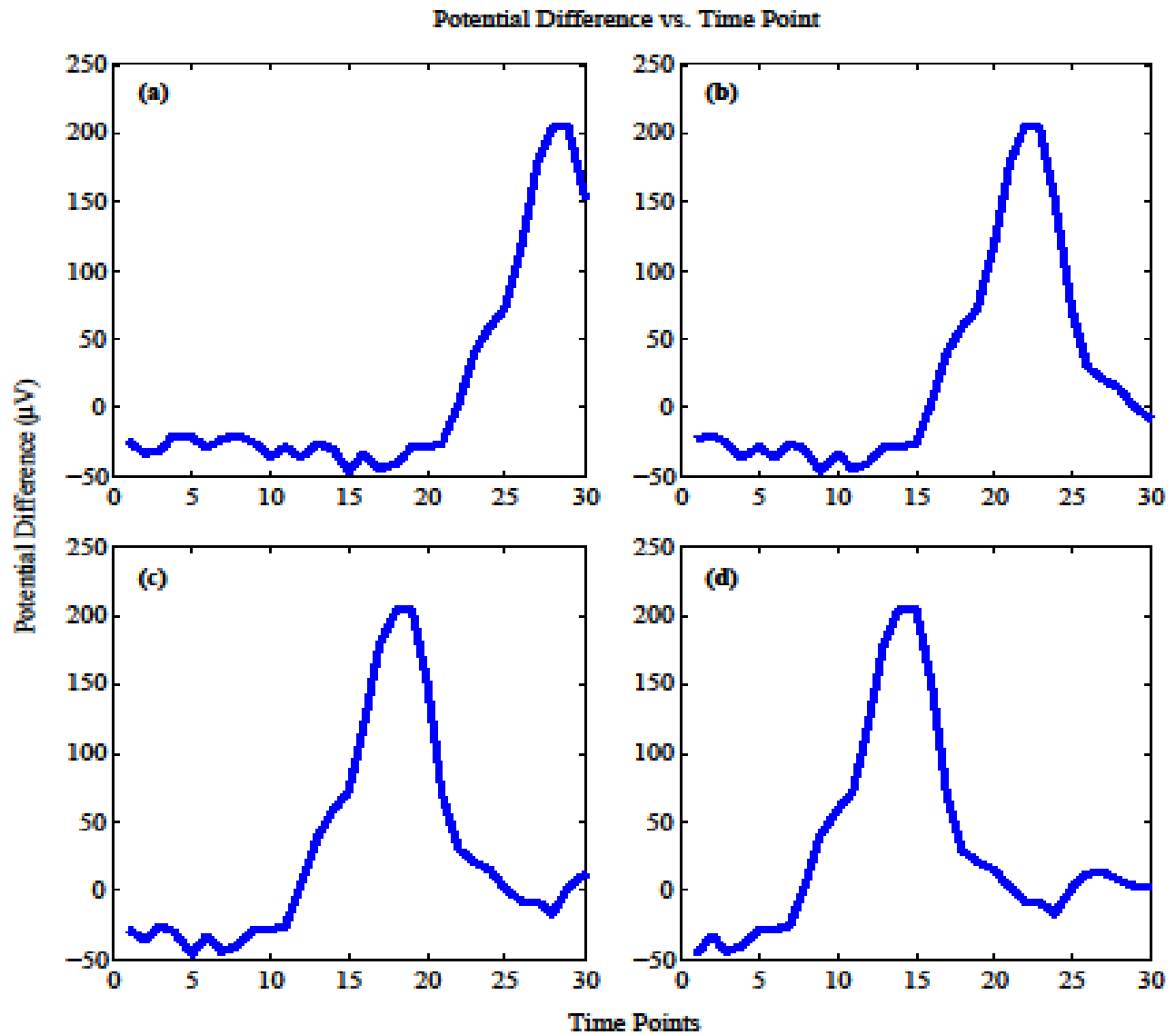
Figure 2.5. Illustration of the shifting by 10 time points showing how a blink from one channel is appearing gradually.

Chapter 3

Results

The simulation of the generalized Ising model as a pattern recognition has been performed using a home made MATLAB program developed by a previous code by Tushar Das. Our initial investigation consisted in testing the convergence of a marginally altered two dimensional pattern to the original pattern. We then created a letter "A" using a 10x10 matrix based on the up spins, and subsequently the corresponding J_{ij}. As we discussed in the methods, another configuration similar to the letter "A" has been also built by mildly perturbing the configuration "A", and we called it "À". What we showed that if we use "À" as an initial configuration, the generalized Ising Model Algorithm will converge to the letter "A" pattern (Figure 3.1).
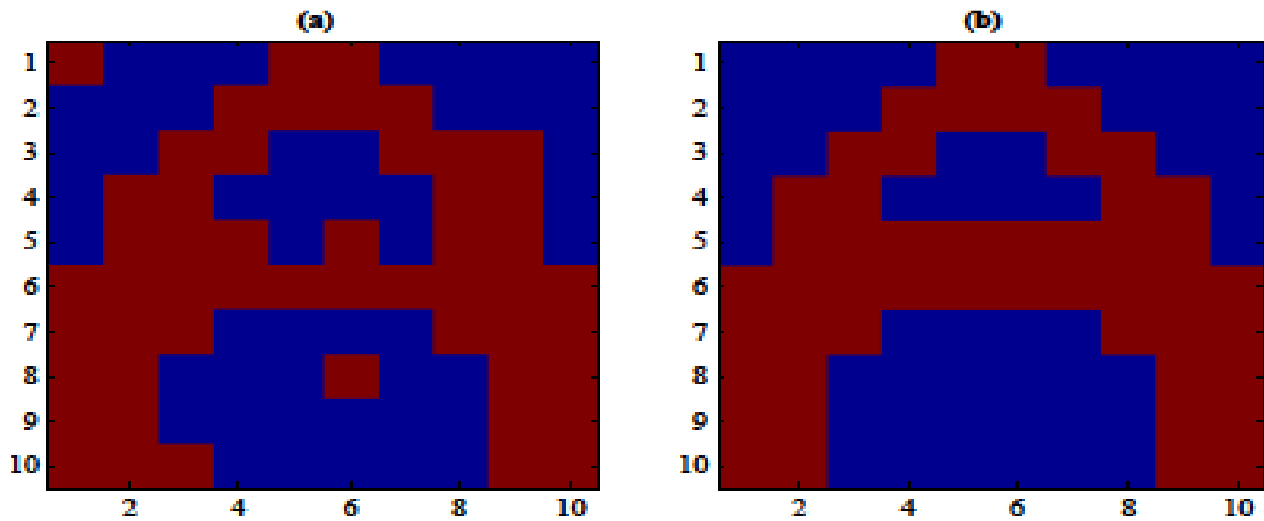
Figure 3.1. Representations of the letter "A" in (a) and the similar configuration "À" in (b). If we start with (a), we will eventually have the pattern in (b).

Similarly, for the 12x30 blink pattern, its J_{ij} has been built, and the 84 blinks extracted from the

second subject were used as an initial configurations and tested for convergence to the average

blink obtained from the very first subject. We found that all the blinks evolved to the blink pattern.

See Figure 3.2 (a) for an example of a blink out of the 84 extracted from the second subject, and

Figure 3.2 (b) for the average blink pattern from the very first subject respectively.
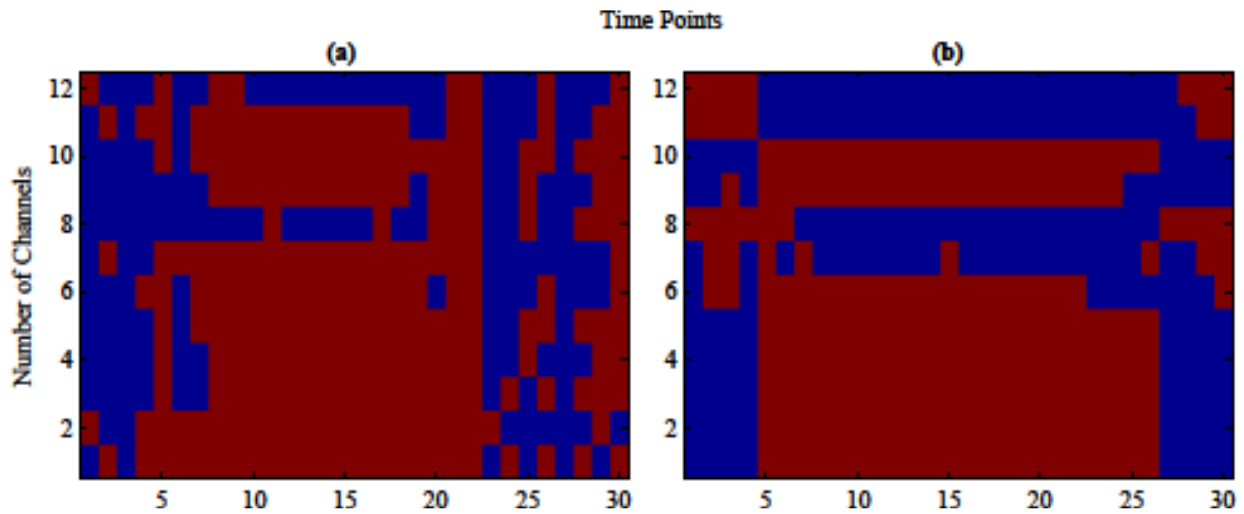
Figure 3.2. Representation of a blink in (a) from the second subject which converged to the
blink pattern in (b) using the generalized Ising Model.

Furthermore, we run a test using different temperatures in order to investigate the convenient

temperature range that could favour blink convergence. Figure 3.3 below illustrates the number

of detected blinks (or which converged to the average blink) from the second subject versus

temperature.
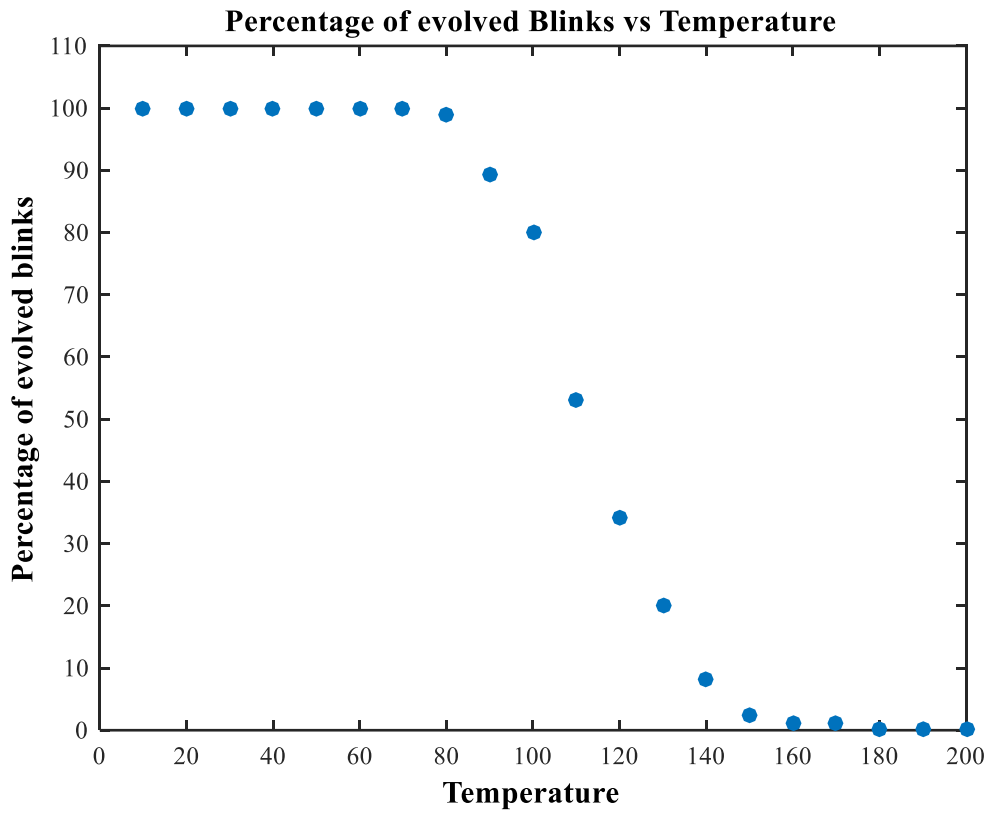
Figure 3.3. Percentage of evolved blinks versus temperature.

We also plotted the thermodynamic properties of the generalized Ising model, like energy, specific heat, magnetization, and susceptibility using a blink as initial configuration. These properties, see Figure 3.4 for their behavior vs temperature, are generally investigated to study the Ising model at criticality.
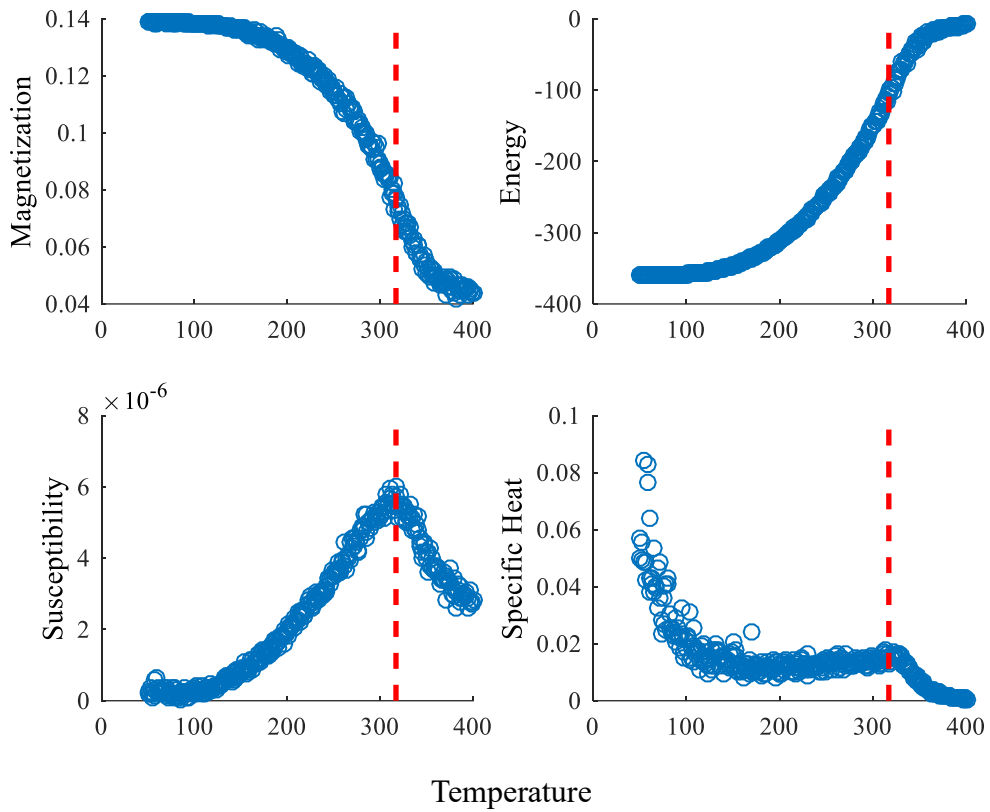
Figure 3.4. Energy, specific heat, magnetization, and susceptibility vs temperature with critical temperature value 317.

As explained in Chapter 2, when the blinks were detected in the EEG data, the epoch of all the 31 channels and 150 time points were substituted by zeros. Figures 3.5 (a) and (b) below illustrates respectively an EEG 10s time window before and after blinks detection.
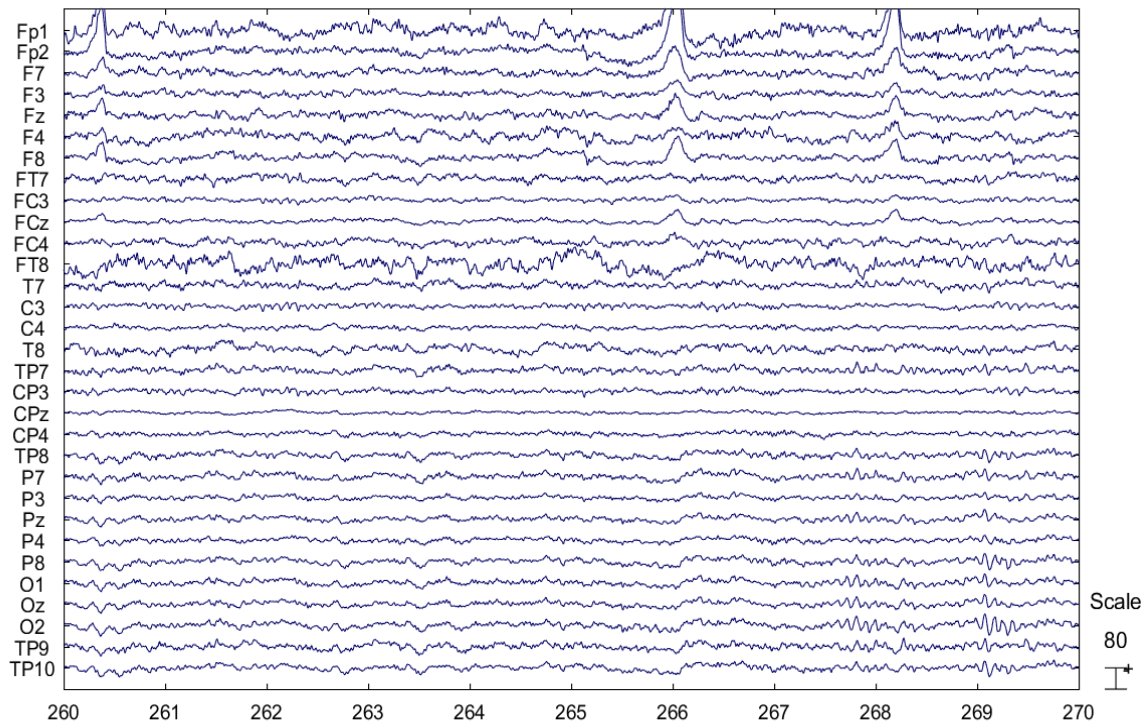
Figure 3.5 (a). EEG recording for 31 channels with three blinks detected at the time 260, 265, and 268 seconds. Voltage is in µV.
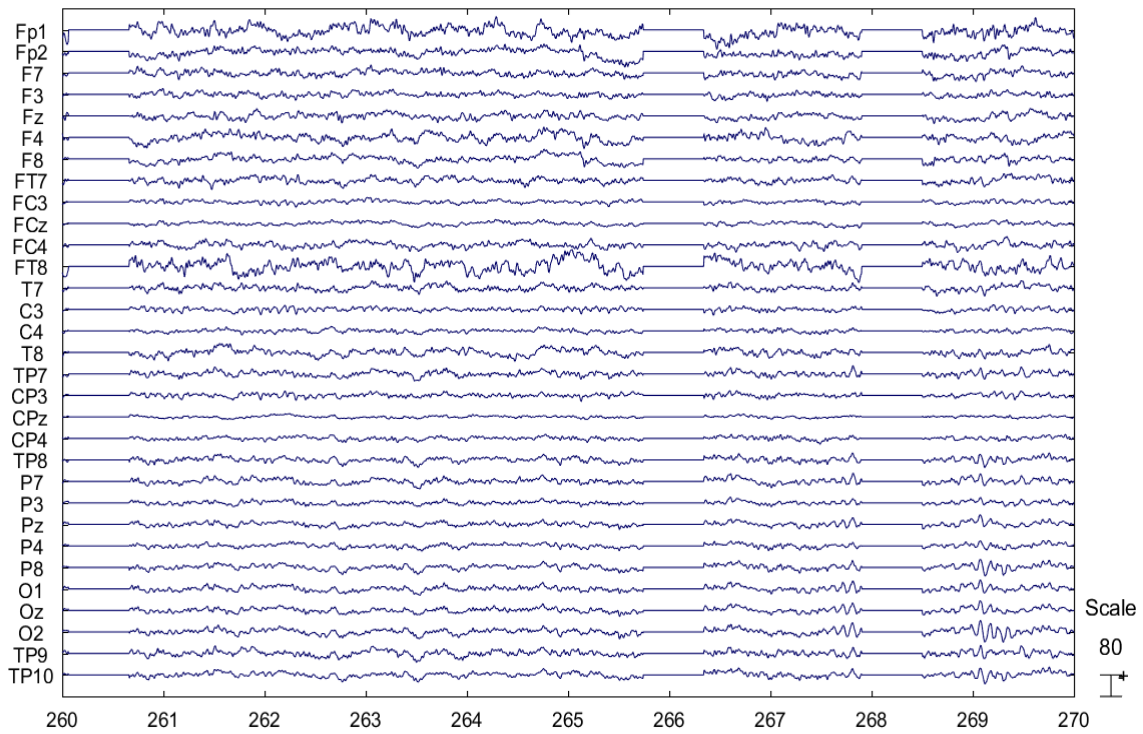
Figure 3.5 (b).EEG recording with three detected blinks appear in the 260, 265, and 268 seconds in 31channels.

In order to quantify the performance of our detection tool and to evaluate the efficiency of the proposed method we calculated respectively the number of True Positive (TP): a blink is recognized as blink, of False Positive (FP): normal activity (not a blink) is recognized as a blink, of False Negative (FN): a blink is not recognized as a blink, and finally of True Negative (TN): normal activity is not recognized as a blink. Then we could estimate respectively the sensitivity, specificity and the accuracy as in Equations 3.1, 3.2 and 3.3.

$$Sen = \frac{TP}{TP+FN} \times 100 \text{ \%} \qquad\qquad (3.1)$$

$$Spe = \frac{TN}{TN+FP} \times 100 \text{ \%} \qquad\qquad (3.2)$$

$$Acc = \frac{TN+TP}{TN+TP+FP+FN} \times 100 \text{ \%} \quad [36] \qquad\qquad (3.3)$$

Table 3.1 provides details of the EEG recordings from six healthy subjects, the number of blinks, the blink frequency, the percentage of TP, TN, FP, FN, Sensitivity, Specificity, and Accuracy in each subject.

Table 3.1. Duration of EEG recording in minutes, number of blinks, blink frequency, percentage of TP, TN, FP, FN, sensitivity, specificity and accuracy for six subjects, together with the mean and standard deviation over the six subjects.

Subjects	Duration (min)	# Blinks	Blink Freq (Hz)	T_p %	F_p %	T_n %	F_n %	Sen. %	Spe. %	Acc. %
Sub1	19.2	165	0.1	9.4	4.2	84.4	2.0	82.4	95.3	93.8
Sub2	17.9	767	0.7	31.1	0.24	54.0	14.6	68.0	99.6	85.0
Sub3	17	225	0.2	7.0	1.5	85.0	6.4	52.4	98.2	92.0
Sub4	14.7	239	0.3	15.6	1.5	81.0	1.8	89.5	98.1	96.6
Sub5	16.3	207	0.2	12.5	11.0	74.0	2.5	83.1	86.7	86.2
Sub6	15.5	390	0.4	19.0	4.0	68.4	8.5	69.5	94.9	87.7
Mean			0.3	16	4	74	6	74	96	90
Std			0.2	9	6	12	5	14	5	5

$(*T_p = \frac{TP}{(TP+TN+FP+FN)} \ 100\% , \ F_p = \frac{FP}{(TP+TN+FP+FN)} \ 100\%, \ T_n = \frac{TN}{(TP+TN+FP+FN)} \ 100\% , \ F_n = \frac{FN}{(TP+TN+FP+FN)}) \ 100\%)$

Chapter 4

Discussion and Conclusion

In this manuscript we presented a new automated method for eye blinking detection based on the generalized Ising model. The Ising model has been implemented to be used as a pattern recognition tool in order to detect blink events in continuous EEG data. One very first ingredient for pattern recognition is the convergence mechanism. Starting from equation (2.2) in Chapter 2, it was shown that the energy for the classical Ising model is minimized when all spins are aligned in the same direction (ordered spins with all +1s or -1s). However, for the generalized Ising model with a particular choice of the couplings as derived from the letter "A" configuration (Figure 2.1, Chapter 2) all the spins are not aligned any longer in the same direction, and the letter "A" configuration has now the lowest energy.

The letter "A" was tested as an example to verify the performance of the Ising model. The idea to extend the procedure to other patterns as well and in particular to blinks has then proven successful. Since the letter "A" has an identified shape, it was very simple to create its pattern. On the other hand, it is challenging to find a single pattern that could identify the blink because of the eye blink variation from subject to subject and at the same subject level itself. We then decided to build the blink pattern by taking the average of the blinks over one single subject for a total of 25 blink events. After creating the blink pattern, we used a second subject to extract independent blink events (84 blinks) and to test the detection method. This way we could have an idea of the detection performance without the bias of the sliding window procedure which we knew necessary anyway

to perform an automatic detection. The results obtained from using the second subject blinks were optimistic as all the blinks in the second subject evolved to the blink pattern at low temperatures. In fact, in order to determine the optimal temperature range to test blink convergence, the number of evolved blinks were plotted vs temperature as presented in Figure 3.3. It became clear then that a temperature less than 80 would have been optimal.

Despite the fact we were not focused in studying the actual phase transition nor the final equilibrium energy, it is beneficial to understand why we should identify the range of temperatures that allow blinks to evolve. To explain more, in thermodynamics, when the temperature increases, it will be followed by an increase of energy and entropy (disordered system) [37] of the spin system. In other words, the entropy at low temperatures can be negligible while at higher temperatures the entropy is high so that the memorized pattern cannot be reached.

The performance of the described methodology was tested by calculating sensitivity, specificity, and accuracy. Given a test sensitivity measures the rate of total true positives, specificity is used to evaluate the possibility of avoiding the false positives, and accuracy measure overall test reliability [36]. The results pointed out that our generalized Ising model based recognition pattern tool offered an accurate blink detection performance, being able to classify the blinks in all six subjects' recordings with a sensitivity of 74 ± 14 %, a specificity of 95 ± 5 % and an accuracy of $90 \pm 5\%$.

In EEG data processing, it is important to isolate and eliminate the EEG contaminations, with as consequence, EEG data loss. Therefore, before introducing an artifacts removal methods, it is important to consider the impact on causing substantial data loss. In our method, the high level of specificity shows that most of the EEG data is preserved. On the contrary due to the important

variation of the blinking behaviour in the different subjects, the sensitivity ended up not being as high as the specificity.

Other approaches such as the Independent Component Analysis (ICA), which is a technique used to linearly isolate mixed signals based on their independent sources, is a widely used Algorithm to remove artifacts in EEG data artifacts. ICA algorithm is provided under EEGLAB toolbox and its decomposed components are called ICs (See Appendix E). However, for researchers who have no previous knowledge or experience in studying EEG artifacts' IC's, it will be challenging to identify the ICs based on their artifacts' topographies.

Using the EOG (Electrooculogram) channels, which are electrodes located above and around the eyes, it is also another well-known technique for eye blinking and movement elimination. This technique implies the recording of the ocular movements. To clean EEG data from such artifacts, EOG data will be regressed out from the EEG data. Nevertheless, this method has its own disadvantages. EOG electrodes can in fact record cerebral signal themselves and that might cause losing a significant relevant part of the neuronal component when performing the regressing out procedure.

In the case of our proposed method, it is simple, and does not require additional electrodes (such as EOG electrodes) and it is not necessary to change the blink pattern for different subjects. Once the blinks are recognized, the data in the blink epoch will be replaced by zeros (or fully removed from the time course). In this thesis we addressed the importance of having blinks-free EEG data. Because of the reoccurrence of blinking artifacts, a number of EEG segments will be affected by these artifacts and removing them manually by visual inspection will be time consuming and require EEG expertise. Therefore, accessible automatic methods can make a significant change in qualifying EEG analysis [25].

4.1 Future Directions

In the previous sections, we discussed the possibility of using the generalized Ising model as a technique to identify eye blink patterns. The introduced model showed optimistic results in recognising eye blinks even if more needs still to be tested in order to improve the accuracy of the pattern recognition procedure. Also despite the method has been implemented for blinks, it could be easily implemented to detect other features in the EEG signal which are not necessarily of artifactual nature. For instance, sleep spindles, vertex waves and k complex patterns, which are very relevant EEG features in stage 2 sleep [38] could be most probably detected with exactly the same procedure developed in this manuscript.
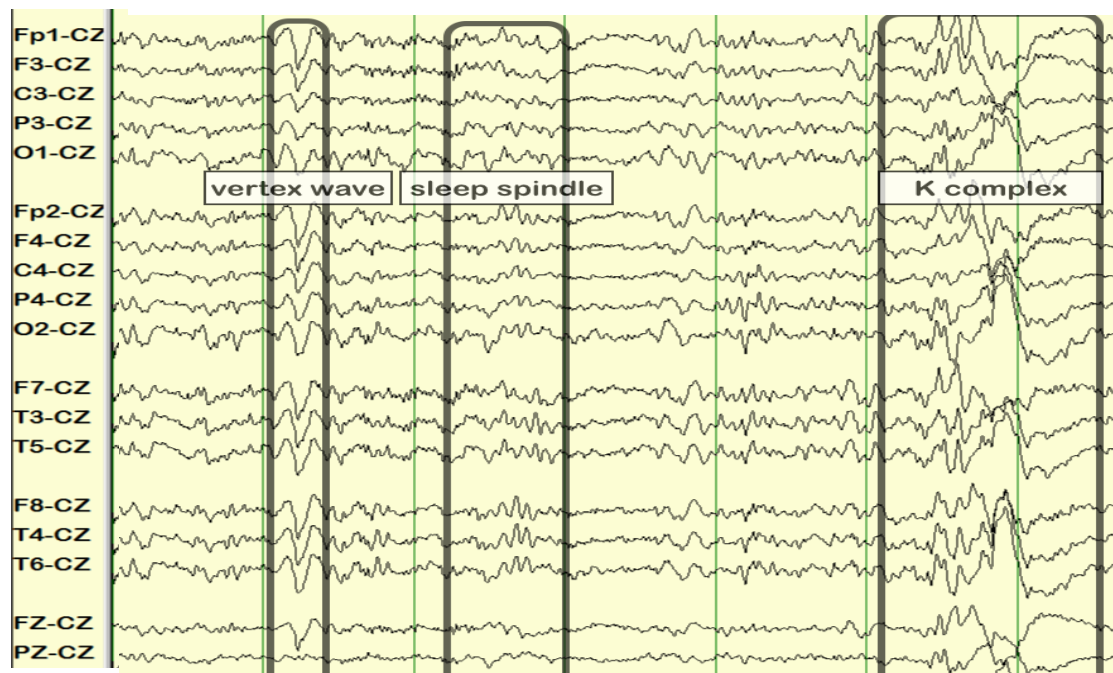
Figure 4.1. Sleep spindle in stage 2 sleep [39].

Bibliography

[1] Teplan, Michal. "Fundamentals of EEG measurement." *Measurement science review* 2.2 (2002): 1-11.

 [2] Lenkov, Dmitry N., et al. "Advantages and limitations of brain imaging methods in the research of absence epilepsy in humans and animal models."*Journal of neuroscience methods* 212.2 (2013): 195-202.

[3] Sur, Shravani, and V. K. Sinha. "Event-related potential: An overview."*Industrial psychiatry journal* 18.1 (2009): 70.

[4] "Neurons." Introduction to Life Science. 2008. Web. 15 Aug. 2016. <http://csls-text2.c.u-tokyo.ac.jp/inactive/06_03.html>.

[5] Nunez, Paul L., and Ramesh Srinivasan. *Electric fields of the brain: the neurophysics of EEG*. Oxford University Press, USA, 2006.

[6] Kirschstein, Timo, and Rüdiger Köhling. "What is the Source of the EEG?."*Clinical EEG and neuroscience* 40.3 (2009): 146-149.

[7] Jurcak, Valer, Daisuke Tsuzuki, and Ippeita Dan. "10/20, 10/10, and 10/5 systems revisited: their validity as relative head-surface-based positioning systems." *Neuroimage* 34.4 (2007): 1600-1611.

[8] Klem, George H., et al. "The ten-twenty electrode system of the International Federation." *Electroencephalogr Clin Neurophysiol* 52.3 (1999).

[9] "Lobes of the brain." Queensland Brain Institute, 2016. Web. 30 Oct. 2016.

<http://www.qbi.uq.edu.au/the-brain/anatomy/brain-lobes>.

[10] EEG. Web. 15 Aug. 2016.

<http://www.medicine.mcgill.ca/physio/vlab/biomed_signals/eeg_n.htm>.

[11] Taywade, S. A., and R. D. Raut. "A review: EEG signal analysis with different methodologies." *Proceedings of the National Conference on Innovative Paradigms in Engineering and Technology (NCIPET'12)*. 2014.

[12] Chabot, Robert J., et al. "Sensitivity and specificity of QEEG in children with attention deficit or specific developmental learning disorders." *CLINICAL ELECTROENCEPHALOGRAPHY-CHICAGO-* 27 (1996): 26-34.

[13] Sanei, Saeid, and Jonathon A. Chambers. *EEG signal processing*. John Wiley & Sons, 2013.

[14] Abo-Zahhad, M., Sabah M. Ahmed, and Sherif N. Abbas. "A new EEG acquisition protocol for biometric identification using eye blinking signals." *International Journal of Intelligent Systems and Applications* 7.6 (2015): 48.

[15] Schomer, Donald L. "The normal EEG in an adult." *The clinical neurophysiology primer*. Humana Press, 2007. 57-71.

[16] Jia, Xiaoxuan, and Adam Kohn. "Gamma rhythms in the brain." *PLoS Biol*9.4 (2011): e1001045.

[17] Herzog, Andrew G. "Catamenial epilepsy: definition, prevalence pathophysiology and treatment." *Seizure* 17.2 (2008): 151-159.

[18] Benbadis, S. R., et al. "Short-term outpatient EEG video with induction in the diagnosis of psychogenic seizures." *Neurology* 63.9 (2004): 1728-1730.

[19] NÚÑEZ, BENITO, and IVÁN MANUEL. *EEG artifact detection*. Diss. 2011.

[20] Park, Hae-Jeong, Do-Un Jeong, and Kwang-Suk Park. "Automated detection and elimination of periodic ECG artifacts in EEG using the energy interval histogram method." *IEEE transactions on Biomedical Engineering* 49.12 (2002): 1526-1533.

[21] Goncharova, Irina I., et al. "EMG contamination of EEG: spectral and topographical characteristics." *Clinical neurophysiology* 114.9 (2003): 1580-1593.

[22] Brunner, Daniel, et al. "Muscle artifacts in the sleep EEG: Automated detection and effect on all-night EEG power spectra." *Journal of sleep research* 5.3 (1996): 155-164.

[23] He, Taigang, Gari Clifford, and Lionel Tarassenko. "Application of independent component analysis in removing artefacts from the electrocardiogram." *Neural Computing & Applications* 15.2 (2006): 105-116.

 [24] Fisch, Bruce."EEG Artifacts.". Web. 15 July. 2016.<https://wiki.umms.med.umich.edu/download/attachments/133925074/EEG+artifacts.pdf?version=1&modificationDate=1392310250000021>.

[25] Ghandeharion, Hosna, and Abbas Erfanian. "A fully automatic ocular artifact suppression from EEG data using higher order statistics: Improved performance by wavelet analysis." *Medical engineering & physics* 32.7 (2010): 720-729.

[26] Tiganj, Zoran, et al. "An algebraic method for eye blink artifacts detection in single channel EEG recordings." *17th International Conference on Biomagnetism Advances in Biomagnetism– Biomag2010*. Springer Berlin Heidelberg, 2010.

[27] Brunner, Clemens, Arnaud Delorme, and Scott Makeig. "Eeglab–an open source matlab toolbox for electrophysiological research." *Biomed Tech* 58 (2013): 1.

[28] Brush, Stephen G. "History of the Lenz-Ising model." *Reviews of modern physics* 39.4 (1967): 883.

[29] Mnyukh, Yuri. "Ferromagnetic state and phase transitions." *arXiv preprint arXiv:1106.3795* (2011).

[30] Niss, Martin. "History of the Lenz-Ising model 1920–1950: from ferromagnetic to cooperative phenomena." *Archive for history of exact sciences* 59.3 (2005): 267-318.

[31] Wu, Fa Yueh, et al. "Comment on a recent conjectured solution of the three-dimensional Ising model." *Philosophical Magazine* 88.26 (2008): 3093-3095.

[32] Cai, Wei. "Introduction to Statistical Mechanics.". Ising Model Hand out. Web. 15 Aug. 2016. <http://micro.stanford.edu/~caiwei/me334/Chap12_Ising_Model_v04.pdf>.

[33] Das, T. K., et al. "Highlighting the structure-function relationship of the brain with the ising model and graph theory." *BioMed research international* 2014 (2014).

[34] Cipra, Barry A. "An introduction to the Ising model." *American Mathematical Monthly* 94.10 (1987): 937-959.

[35] Binder, Kurt. "Introduction: Theory and "technical" aspects of Monte Carlo simulations." *Monte Carlo Methods in Statistical Physics*. Springer Berlin Heidelberg, 1986. 1-45.

[36] Zhu, Wen, Nancy Zeng, and Ning Wang. "Sensitivity, specificity, accuracy, associated confidence interval and ROC analysis with practical SAS® implementations." *NESUG proceedings: health care and life sciences, Baltimore, Maryland* (2010): 1-9.].

[37] Kardar, Mehran. *Statistical physics of particles*. Cambridge University Press, 2007, Page 13.

[38] De Gennaro, Luigi, and Michele Ferrara. "Sleep spindles: an overview."*Sleep medicine reviews* 7.5 (2003): 423-440.

[39] Strayhorn, David. *EEG Atlas*. Web. 15 Aug. 2016. <http://eegatlas-online.com/index.php/en/alphabetical-index/vertex-waves-guest>.

Appendices

1. Appendix A: Generalized Ising Model Algorithm

```matlab
%================================================================
=========================================

% Blink Detection in Spontaneous EEG using the Generalized Ising Model

% and Monte Carlo simulation (modified from Tushar K. Das Algorithm)

%================================================================
=========================================

clear all

load blink12by30pat_d.mat

load('J_blink12.mat')

load('eeg_data.mat')%  eeg data

no_slidingpoints=10;

[L1,L2]=size(eeg); % L1,L2 dimensions of the eeg data

T=5; % Temprature

no_slices= floor(((L2-150)/no_slidingpoints)+1);% each slice is 12*150 matrix

J=J_blink12; % (360*360) matrix, this number came from considering the blink is 12*150

matrix,

%and J is describing how all these pixels connecting to each other,so J is 360*360 matrix

N=360; % dimension of J matrix

blinkpat=blink12by30pat_d;% the blink pattern

no_flip=N*5;

randcoord = randperm(N);

eegdata=zeros(12,150,no_slices);
```

```matlab
x=zeros(1,no_slices);

eeg2=eeg_data(1:12,:);

tic

for nnn=1:no_slices

    % Running through the EEG data by taking only the 12 frontal and

    % prefrontal (Fp1, Fp2 ,F7,F3 ,Fz ,F4 ,F8 ,FT7 ,Fc3 ,Fcz ,Fc4 ,FT8 ) channels, then

    % taking 12*150 epoch each time from the data by sliding by 10 and the

    % digitize it make it ( +1 and -1)

    F=eeg2(:,1+(nnn-1)*no_slidingpoints:150+(nnn-1)*no_slidingpoints);%

    eegdata(:,:,nnn)=F;

    for iii=1:12

        for jjj=1:30

            % Downsamling from 150 time points to 30

            %by taking the mean of every 5 points

            eegdatared30(iii,jjj,nnn)=mean(eegdata(iii,1+(jjj-1)*5:5+(jjj-1)*5,nnn));

            %digitizing based on a baseline=14

            if   eegdata(iii,jjj,nnn)>=14

                eegdatared30_d(iii,jjj,nnn)=1;

            elseif   eegdata(iii,jjj,nnn)<14

                eegdatared30_d(iii,jjj,nnn)=-1;

            end

        end

    end

end
```

```matlab
end

for mm=1:no_slices

    spv=eegdatared30_d(:,:,mm);


    nn=0;
    for i = 1:no_flip


        nn = nn + 1;
        %        if (nn >N);
        %            nn = nn-N; % the nn will not go over N
        %             randcoord = randperm(N); %  rerandomise again
        %        end
        Flip = randcoord(nn);% random choice
        % Compute the change in energy
        dE=0;
        for j=1:N
            if (j~= Flip)
                dE = dE +J(Flip,j)*spv(j);
            end
        end
        dE=2*dE*spv(Flip);
        if (dE <= 0)
            spv(Flip) = - spv(Flip);
```

```matlab
        elseif (rand <= exp(-dE/T))

            spv(Flip) = - spv(Flip);

        end



    end

    ener = 0;

    for ii = 1:N;

        for jj=1:N

            ener = ener - J(ii,jj)*spv(ii)*spv(jj);



        end

    end



    if spv==blinkpat; % comparing  whether the epoch evolved to the blink pattern or not



        x(mm)=mm;

    end

end

y=find(x);%  this matrix represents the number of slices that evolved to blink



% replacing the epochs that evolved to blink with zeros as a mark that of containing a blink

slices_index=zeros(no_slices,150);

no_slidingpoints=10;
```

```matlab
data=eeg;

clean_eegdata=eeg;


for i=1:no_slices

    slices_index(i,:)=1+(i-1)*no_slidingpoints:150+(i-1)*no_slidingpoints;

end

L=length(y);

for m=1:L-5


    if y(m+3)==y(m)+3

        c=max(data(1,slices_index(y(m)):slices_index(y(m)+4,150)));

        c1=find(data(1,slices_index(y(m)):slices_index(y(m)+4,150))==c);

        c2=slices_index(y(m))+c1-1;


        clean_eegdata(:,c2-74:c2+75)=0;


    end

end

toc

clear Flip J N T  count  i ii j jj no_flip randcoord  temp nn dE baseline m mm  iii jjj nnn F ener
eegdata data data2 dE c c1 c2 blinkpat blink12by30pat_d eeg eegdata eegdata30 eegdata30_d
ener F  J_blink12 L1 L2  no_slices no_slidingpoints slices_index spv x m y eegdatared30
eegdatared30_d L eeg2
```

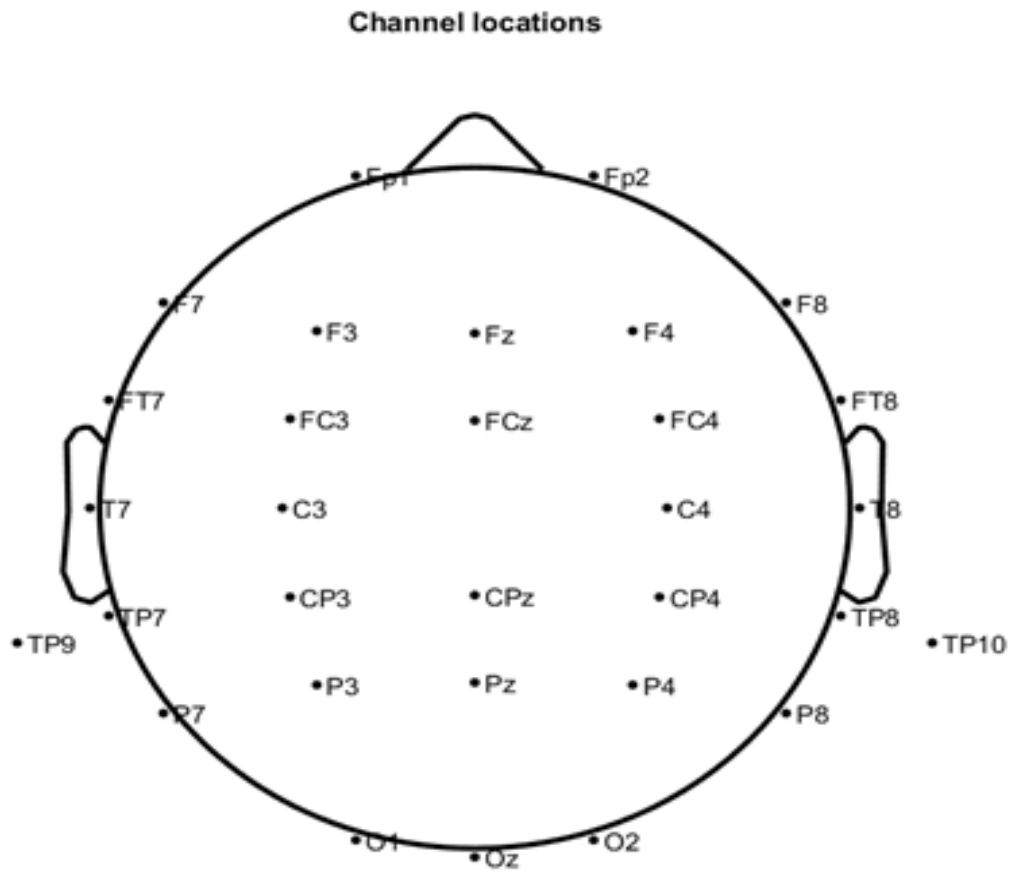## 2. Appendix B: 2D Channel Locations.

Channel locations

Figure 2. 2D channel locations plotted using EEGLAB.

3. Appendix C: Gender, Age, and Types of EEG Data of the Subjects Used in This Research.

Subjects	Gender	Age	EEG Data
S1	Female	48	Blinks Only
S2	Female	69	Blinks Only
S3	Male	32	EEG 32 channels
S4	Male	31	EEG 32 channels
S5	Male	30	EEG 32 channels
S6	Male	62	EEG 32 channels
S7	Male	52	EEG 32 channels
S8	Female	44	EEG 32 channels

Table 2. Representation of the gender, age, and types of EEG data from the eight subjects that used in this study.
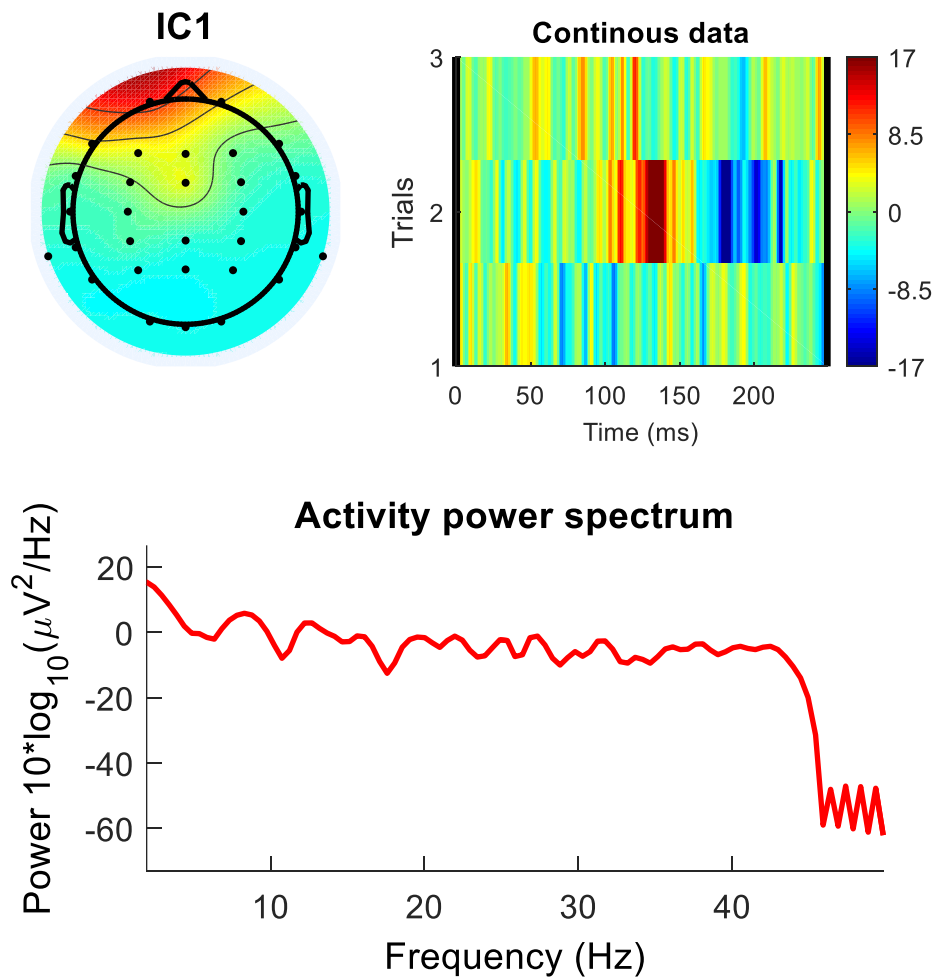
4. Appendix D: Blink Topography

Figure 3. Representations of a blink by different graphical system. Starting from the upright image, it represents the IC blink topography, on the left graph represents the image of the blink and finally the bottom plot of the power spectrum vs frequency.

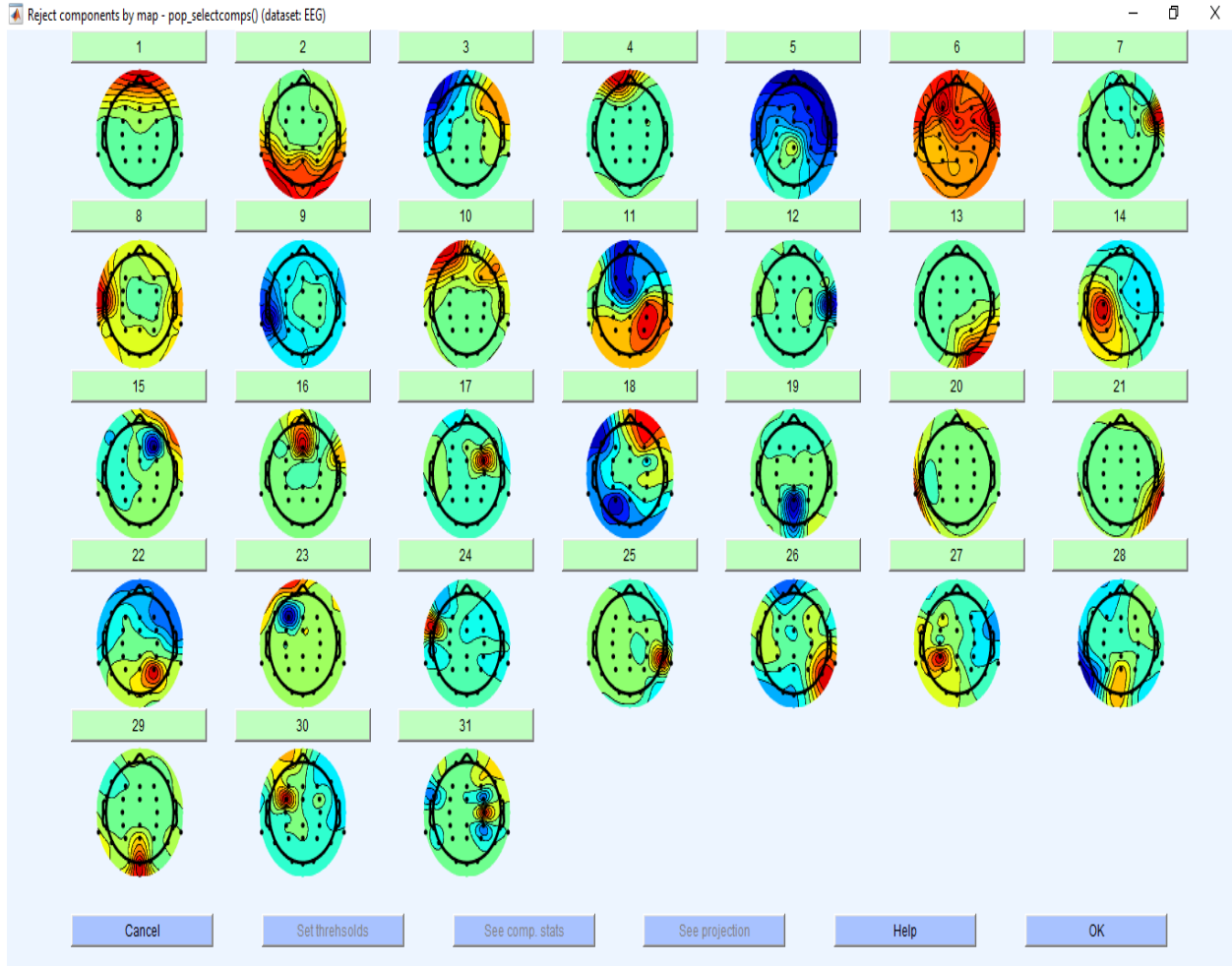# 5. Appendix E: Independent Component Analysis Topographies

Figure 4. 31 ICs for the third subject. Once the component recognized as an artifact, it can be rejected using EEGLAB.

Curriculum Vitae

Name: Marwa Dawaga

Post-Secondary University of Tripoli, Tripoli
Education and B.Sc. degree in Physics
Degrees: 2003-2007

Honors and   Teaching Assistant in Physics Department, Tripoli University
Awards: Ministry of higher education and research scholarship, Libya

Conference   Poster Presentation Fallona Family
and Presentation: Interdisciplinary Showcase 2015