Advanced Statistical Machine Learning Methods for the Analysis of Neurophysiologic Data with Medical Application

by

Julià Lluís Amengual Roig

Departament de Llenguatges i Sistemes Informàtics
UNIVERSITAT POLITÈCNICA DE CATALUNYA

A thesis submitted in partial fulfillment for the degree of Master of Science

June 2010

Advanced Statistical Machine Learning Methods for the Analysis of Neurophysiologic Data with Medical Application

Julià Lluís Amengual Roig

Thesis Summary

Transcranial magnetic stimulation procedures use a magnetic field to carry a short-lasting electrical current pulse into the brain, where it stimulates neurons, particularly in superficial regions of the cerebral cortex. It is a powerfull tool to calculate several parameters related to the intracortical excitability and inhibition of the motor cortex. The cortical silent period (CSP), evoked by magnetic stimulation, corresponds to the suppression of muscle activity for a short period after a muscle response to a magnetic stimulation. The duration of the CSP is paramount to assess intracortical inhibition, and it is known to be correlated with the prognosis of stroke patients' motor ability. Current mechanisms to estimate the duration of the CSP are mostly based on the analysis of raw electromyographical (EMG) signal and they are very sensitive to the presence of noise.

This master thesis is devoted to the analysis of the EMG signal of stroke patients under rehabilitation. The use of advanced statistical machine learning techniques that behave robustly in the presence of noise for this analysis allows us to accurately estimate signal parameters such as the CSP. The research reported in this thesis provides us with a first evidence about their applicability in other areas of neuroscience.

Keywords: Electromiography, Stroke ,Variational Bayesian Generative Topographic Mapping, Index of Variability, Silent Period.

Acknowledgements

I would like to thank Dr. Alfredo Vellido, Dr. Ivan Olier and Dr. Antoni Rodríguez-Fornells for them guidence, help and supporting in this work. Them supervision has made the study and redaction of this thesis a very rewarding experience, with many challenges and hard work.

I am very gratefull to Dr. Carles Grau for his complete trust on me since I am working in Psychiatry and Clinical Psicobiology Dept of the University of Barcelona.

I would also like to thank the people of the *Cognition and Brain Plasticity Unit*, in Hospital de Bellvitge, especially Nuria Rojo for her help and participation in our project research with stroke patients. My gratitude extends to Nóra Schulcz, for her personal supporting during the redaction of this master thesis.

I acknowledge the economic support of *La Marató de TV3* in our project *Musical Supported Therapy Induces Plasticity on Sensorimotor Cortex after Stroke*

Last, but by no means least, I want to thank those people who are able to think that the power of research advance resides on divergence.

Contents

List of Figures

Dedicated to my grandfather, the man who always trusted on me.

Chapter 1

Introduction

1.1 Introduction and Motivation

Recent years have witnessed an unprecedented assault on one of the most ambitious goals of contemporary science: the evidence-based investigation of the workings of large-scale natural neural networks, or, in other words, the workings of of the human brain.

Computational neuroscience has become on the key columns of this intensely multi-disciplinary research effort that gathers its strength from fields as diverse as systems biology, genetics, bioinformatics, signal processing, artificial intelligence, and psychology.

The area of cognitive neuroscience is benefiting from the use of the most technologically advanced tools developed within each of these fields, with the purpose of unraveling high-level cognitive processes. The data acquired with some of these tools has become complex enough as to require advanced data analysis techniques with capabilities beyond those of traditional statistics.

Physiological data usually have a strong component of noise that must be processed using robust procedures in order to extract usable knowledge from them. This is the case both for the purpose of research and for routine medical practice. In electromiographical (EMG) recordings, the electrical activity of motor neurons is recorded from the skin surface above the muscle, and not at the muscle itself, in order to avoid unnecessary discomfort in the analyzed subject. Also, in electroencephalography data recordings (EEG), external electrodes are used to register the electrical activity of the brain from the skin of the scalp, due the obvious difficulties that an intracranial recording would imply.

The amount of noise in the recorded data is not only related to the limitations of the data acquisition techniques, but also, importantly, to the conditions in which these data are acquired. Neurophysiological data are very sensible to motion. EEG recordings, for instance, are sensible to eye blinking, causing large artifacts on the register and rendering the data useless in that time-window. EMG recordings are very sensible to muscle activation, and it is very important to maintain the muscle in a state of relaxation during data collection.

In clinical research, the majority of subjects participating in studies are patients affected by a given pathology. This often makes data acquisition, processing, and analysis rather difficult undertakings. For the reasons outlined above, physiological data in general and EMG data,

which are the concern of this thesis, in particular, would benefit from the development and use of data analysis models that behaved robustly in the presence of noise.

This thesis contributes in a sensitive area of neuroscience, in the frontier with clinical medicine: the rehabilitation of patients affected by stroke. A brain stroke is the rapidly developing loss of brain functions due to disturbance in the blood supply to the brain. This may cause the large-scale death of neuron populations located in the area where the blood disturbance affects.

Depending on this location, different symptoms and conditions will occur. For example, if stroke affects occipital region of the brain, there is a high probability of loss of visual skills. Because of the topology of the pyramidal tract (fibers of neurons related with motor skills), the majority of stroke patients suffer motor disabilities on the body side contralateral to the lesion side on the brain. The improvement of such skills depends on the rehabilitation performed in the acute phase of stroke (under 6 months after stroke).

In this period, many brain mechanisms are activated to replace the loss of the affected functions, such as motor impairments. The onset of these different mechanisms is called *plasticity*. Basically, plasticity is the ability of the brain to change, modify and modulate itself structurally and functionally, in order to adapt to contextual novelties. Many studies of plasticity in stroke patients show that there exist correlations between brain parameters and the prognosis of the patient (i.e., the prediction of the function recovery after acute stroke phase). These parameters are related to neuropsychological tests and brain stimulation results.

Transcranial Magnetic Stimulation (TMS) is a non-invasive method used to excite the elementary unit of the nervous system: neurons. Using a coil attached to special capacitor, weak electric currents are induced in the tissue by rapidly changing magnetic fields (electromagnetic induction). This way, brain activity can be triggered with minimal discomfort, and the functionality of the circuitry and connectivity of the brain can be studied. Many applications of cortical stimulation has been explored. For example, TMS was applied to understand how the visual cortex works in blind people in relation with memory, demonstrating that this visual cortex has an important role in working memory processing, as compared with non-blind subjects [70]. TMS has also been applied to patients affected by depression. It has been demonstrated that repetitive application of magnetic pulses on frontal lobe regions leads to an important improvement in the mood of patients with anxiety and emotional disorders [58]

This technique has its more obvious application in motor cortex analysis. The motor cortex has been studied in depth with brain stimulation (over 3,200 published articles since the 80's in U.S. National Library of Medicine, NIH). For a large part of these publications, The EMG recording was used to register the signal of the corresponding muscle activation. Through this signal, many excitatory (like the amplitude of the signal muscle potential) and inhibitory parameters can be studied.

One of the most important inhibitory parameters is the Cortical Silent Period (CSP) [65]. The CSP is a refractory period in the EMG signal after motor cortex stimulation with voluntary preactivation of the target muscle, showing both cortical and spinal cord inhibition. The duration of this de-activation interval is an important parameter in neurophysiology. More specifically, in stroke rehabilitation, many studies have demonstrated that the CSP shortens during the recovery of the affected limbs, being a reliable indicator of therapeutic progress [46, 47]. EMG recordings of stroke patients are often difficult in both acute and chronic patients. In particular, the EMG signal recording from stroke patients is fraught with difficulties that become explicit in the form of a CSP measurement protocol. In this protocol, it is critical to ensure that the patient is able to sustain a stable voluntary contraction of the muscle, even when the TMS pulse is performed,

in order to obtain a reliable EMG post-CSP contraction signal and, with this, the offset of the silent period. Some stroke patients have great difficulty in maintaining the muscle contraction in any stable way. The result is that some spurious low-amplitude EMG signal appears during the CSP.

The existing CSP measuring methods in common use are yet imprecise and known to yield a significant error due to their sensitivity to noise, which is commonplace in this kind of data. This is a call for the development of EMG signal analysis methods that behaved robustly in the presence of these data and, specifically, methods for the robust estimation of the CSP in EMG recordings. The current thesis is a response to such call.

For this, we resort in this thesis to a manifold-constrained Hidden Markov Model, which formulation within a variational Bayesian framework imbues it with regularization properties that minimize the negative effect of the presence of noise in the EMG multivariate times series. A novel index of variability is defined for this model, and it is shown that it is capable to estimate the duration of the CSP by accurately pinpointing its offset time.

The remaining of this master thesis is structured into five chapters, as follows:

- **Chapter 1** It describes the principles and applications of Transcranial Magnetic Stimulation, through a historical and technical approach. It also introduces the concept of induced inhibition with magnetic stimulation and provides a precise description of the silent period.
- **Chapter 2** This chapter aims to summarily introduce the a medical background and theoretical principles of brain stroke, describing typology, symptoms and conditions. The biological basis and principles of neuro-rehabilitation and plasticity are introduced, which are basic to understand the improvement mechanisms of the brain.
- **Chapter 3** It provides a description of the machine learning technique proposed for the robust estimation of the CSP duration. The Generative Topographic Mapping (GTM) technique, and its principled extension, Generative Topographic Map Through Time (GTM-TT) are described. The latter is suited to deal with the unsupervised analysis of multivariate time series.
- **Chapter 4** The theoretical approach described in chapter 3 is used in a battery of experiments with artificial data and real EMG signal corresponding to human controls and stroke patients.
- **Chapter 5** concludes the master thesis with a summary of its main contributions. Furthermore, a discussion on future directions, as well as of open research questions, is summarily outlined.

Chapter 2

Medical Background

2.1 Medical Background: Stroke and Neorological Basis of Rehabilitation

2.1.1 Brain Stroke

A stroke is the rapidly developing loss of brain functions due to disturbance in the blood supply to the brain. This can be due to *ischemia*, which is a lack of glucose and oxigen supply caused by thrombosis or embolism (Fig. 2.1), or due to a *hemorrhage*, which is a huge loss of blood because of a stroke in a vessel. In any case, as a result, the affected area of the brain is unable to function, leading to inability to move one or more limbs on one side of the body; cognitive impairing such inability to understand or formulate speech (*aphasia*); or inability to perceive one side of the visual field (*homonymous hemianopsia*).

Stroke can cause permanent neurological damage, complications and even death. It is in fact the leading cause of adult disability in the developed countries. In an ischemic stroke, blood supply to part of the brain is decreased, leading to dysfunction of the brain tissue in that area. There are three reasons why this might happen:

- Thrombosis, which is the obstruction of a blood vessel by a blood clod forming locally (Fig 2.2).
- Embolism, which is the obstruction due to an embolus from elsewhere in the body.
- Systemic hypoperfusion, a generic decrease in blood supply.

Usually, symptoms occur suddenly and are often most severe a few minutes after they start, because most ischemic strokes begin suddenly, develop rapidly, and cause death of brain tissue within minutes to hours [48]. After the initial period, most strokes become stable, causing little or no further damage. Strokes that remain stable for 2 to 3 days are called completed strokes. Sudden blockage by an embolus is most likely to cause this kind of stroke.

Many different symptoms can occur, depending to a great extent on which artery is blocked and, thus, on which part of the brain is deprived of blood and oxygen. For example, damage to the frontal lobes causes loss of the ability to solve problems, plan actions and the fluency of speech

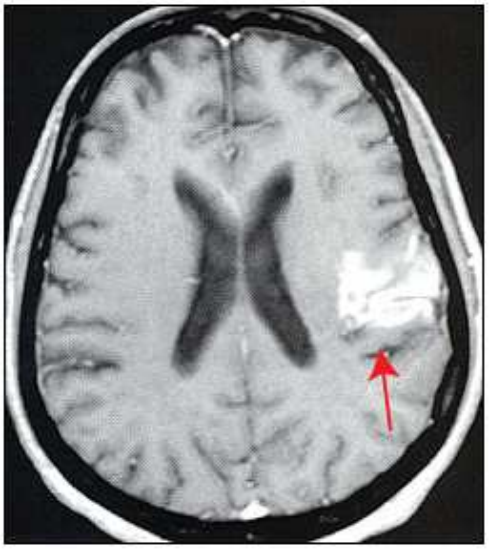

FIGURE 2.1: Structural magnetic resonance image of a ischemic stroke (arrow), located on right tempo-parietal area.

is significantly reduced. Also, fronto-central damage causes loss of the motor ability and control of the contralesional side of the body, producing a level of hemiparesis which depends on the lesion.

Prognosis, a medical term to describe the likely outcome of an illness, is a very important issue in the treatment of stroke patients. About 10% of people who have an ischemic stroke recover almost all normal function, and about 25% recover most of it. About 40% of people have moderate to severe impairments requiring special care, and about 10% require care in a nursing home or other long-term care facility. During the first few days after an ischemic stroke, doctors usually cannot predict whether a person will improve or worsen. Younger people and people who start improving quickly are likely to recover more fully. About 50% of people with one-sided paralysis and most of those with less severe symptoms recover some function by the time they leave the hospital, and they can eventually take care of their basic needs. They can think clearly and walk adequately, although use of the affected arm or leg may be limited. Use of an arm is more often limited than use of a leg. Most of the impairments that remain after 12 months become permanent.

2.1.2 Rehabilitation after Stroke

Disability associated with hemiplegia of hemiparesis markedly limits independent living and social participation in at lest half of all stroke survivors [24]. Reduced levels of exercise and daily activity as a consequence of disability can increase risk factors for recurrent stroke and associated cardiovascular disease.

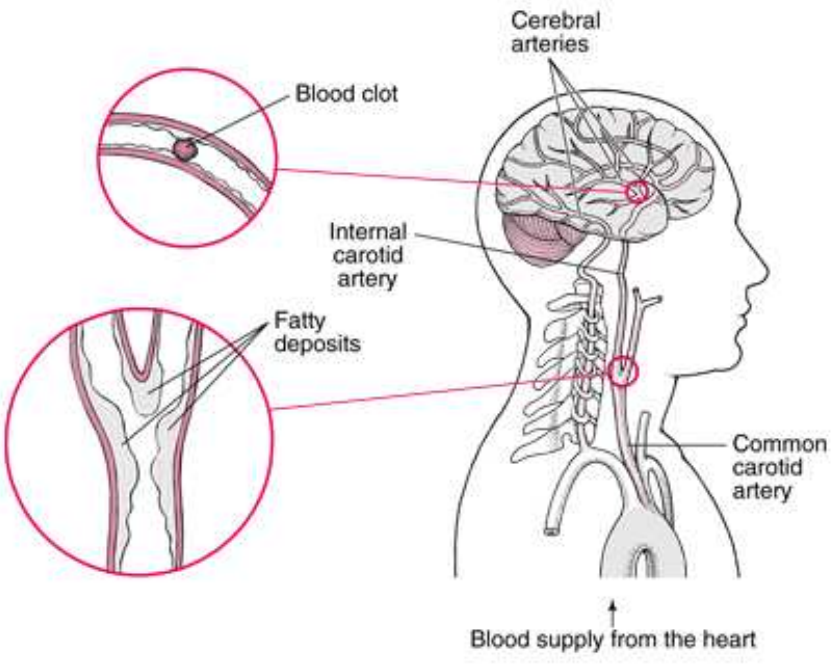

FIGURE 2.2: Anatomical diagram of stroke. Thrombosis, the obstruction of a blood vessed by blood clod, it is one of the most common causes of brain stroke.

Within several days after the onset of a stroke, clinicians can begin to promote functional recovery in their patients. It should be possible to tap into fundamental cellular and molecular events associated with injury in an attempt to lessen impairments, such as the weakness and loss of coordination resulting from hemiparesis, and disabilities, such as limitations in the use of the affected upper extremity for self care, or slow and unsafe ambulation.

In this section, we will describe the fundamental basis of neuro-rehabilitation and some examples of task-oriented practice to improve motor skills.

2.1.2.1 Basic mechanisms of neuroplasticity

Initial motor gains after stroke might result from the resolution of reversible injuries to neurons and glia, such as alterations in membrane potentials, axon conduction or neurotransmission. Reorganization of spared assemblies of neurons that represent motor actions within the sensorimotor cortex, as well as in transcortical, ascending and descending pathways, seems to accompany further improvements in motor skills [24].

Contributions from more widely distributed cortical and subcortical regions, including cerebral systems for perception, attention, motivation, executive planning, working memory, as well as explicit and implicit learning, might be required to compensate for strategies that the injured brain can no longer support. The ability to strengthen muscles and reach an appropriate level of cardiovascular fitness also depends on these non-motor systems.

Adaptability of the building blocks of the movement

The skeletomotor system in humans includes at least seven separate descending corticospinal tract pathways. These tracts seem to rely heavily on intrinsic pattern generators and motor primitives organized within the motor pools of the spinal cord [14]. Spinal systems provide automatic movements across joints alternating flexion and extension of the limbs of stepping. Lesions on these pathways cause loss of movement patterns, along with weakness, fatigability when performing repetitive movements, specially of the elbow, wrist and fingers for reaching and grasping.

Collections of neurons that innervate the spinal motor neurons of a singe muscle intermingle with those that innervate other muscle across one or more joins. This organization allows neurons that innervate different muscle to learn new and fine movement pattern across together, bound by Hebbian long-term potentiation, and other mechanisms that determine synaptic efficacy [49]. This redundancy created by overlapping supra-spinal descending projections can be used to re-learn a movement pattern when some of primary motor cortex or the cortico-spinal pathway has been infarcted. If enough of their pathways have been spared, other parts of the distributed cortical and subcortical sensorimotor system can be recruited to drive the motor pools in the spinal cord. Moreover, uncrossed fibers of the corticospinal tract and axons that recross might provide input that can partially compensate for the loss of motor pathway function.

If an adequate residual percentage of ascending sensory and descending motor contributions are present, carefully chosen pro-active paradigms might enable adequate reaching and grasping with the hand, and walking.

Biological changes

In some animal models of stroke, M1 and related motor cortices and the spinal cord evolve robust changes in their structure and function in response to specific types of motor training. Skills training induces synaptogenesis (creation of new synaptic connections), synaptic potentiation and reorganization of movement representations within the motor cortex. This plasticity supports the production and refinement of skilled movement sequences. Strength training, by contrast, can alter the excitability of spinal motor neurons and induce synaptogenesis within the spinal cord, but does not alter the organization of the motor map, i.e., can induce subcortical but not cortical changes [1]

Experimental studies indicate that the biological changes associated with practice-induced plasticity are molecular (e.g.,changes in gene transcription, protein regulation or neurotrasmitter release), morphological (e.g, growth of dentritic spines associated with long-term potentiation at synapses), and physiological (e.g., excitation and inhibition among assemblies of neurons). It has been proved in caged rodents that exercise leads to changes in gene expression, including upregulation or downregulation of genes encoding molecules associated with learning and memory, and also to neurogenesis. [72]

The biological responses to exercise in patients might be less pronounced than those in relatively experience-deprived, genetically homogeneous laboratory animals [25]. Responses are also likely to depend on how long after stroke the exercise is initiated, the amount of exercise administered, and the duration and type of the task practiced by a patients.

2.1.2.2 Task-oriented practice for skills improvement: Music in rehabilitation

In terms of improving daily functioning, task-specific training seems to benefit stroke patients more than general exercise does (as it does in healthy subjects who wish to learn a new motor skill). One of the problems in demonstrating the specific effects in practice of any given task across rehabilitation trials has been the low intensity of training, which might limit the robustness of outcomes [43]. In addition, responsiveness to training has been observed mostly in patients who have retained reasonable motor control, such as being able to at least partially extend the wrist and fingers or flex the hip and extend the knee of the hemiparetic side.

However, effectiveness of classical approaches in rehabilitation methods, such as rehabilitation based in repetitive manipulation of objects and movement training of the affected side, has been found to be quite limited[63]. As a result, a need for efficient motor rehabilitation approaches still remains.

Relatively new concepts and knowledge about plasticity have become the key for this research on new rehabilitation methods. For example, constraint-induced therapy, which consist on the use of the impaired extremity while immobilizing the healthy extremity for several hours per day, has been shown to lead to functional reorganization. This has been demonstrated by TMS and PET (Positron Emision Tomography). Further studies show that plastic reorganization of neuronal networks may play an important role in recovery after brain injuries, ischemic lesions after stroke, or degeneration processes.

In this field, animal studies have shown that cortical plasticity is increased by the behavioral relevance of the stimulation and its motivational value. Previous studies have shown very rapid plastic adaptation due to music performance which is not restricted to cortical motor areas but also involves auditory and integrative auditory-sensorimotor circuits [51]. Also, because of the hight motivational value of music, and in light of the aforementioned studies on auditory-sensiromotor coupling, many studies are designed entailing active music made in the rehabilitation of stroke patients [60].

Strokes are known to lead to an impairment of propioceptive feedback information, and it has been suggested in many studies that propioceptive reafferences play an important role in updating the internal representations during movement. Other studies have suggested that movement errors committed by moderately to severely hemiparetic patients were due to impaired interpretation of propioceptive information concerning limb position [2, 3]. Therefore, auditory feedback, may serve to counteract this deficit, while standard occupational and physiotherapy approaches use either unspecific sensory stimulation that does not require focussed attention or discriminative effort, or provide only global feedback as to whether or not a movement is achieved.

2.1.3 Transcranial Magnetic Stimulation and Stroke

Ischemic Stroke frequently leads to impairment of upper limb motor function, after which a variable degree of motor recovery is seen [71]. Functional Imaging in humans and physiologic observations in animal models [37] suggest than recovery of function is associated with extensive reorganization of the motor system at the cortical level, presumably to maximize control of remaining motor output. TMS has also been used in human stroke patients to probe corticospinal and intra-cortical physiology. For example, reduced corticospinal excitability from the affected hemisphere reflects damage to the corticospinal connection [16] (See Fig.2.4); whereas

increases in intra-cortical excitability on both hemispheres [45] reflect changes in intrinsic circuits of the cortex.

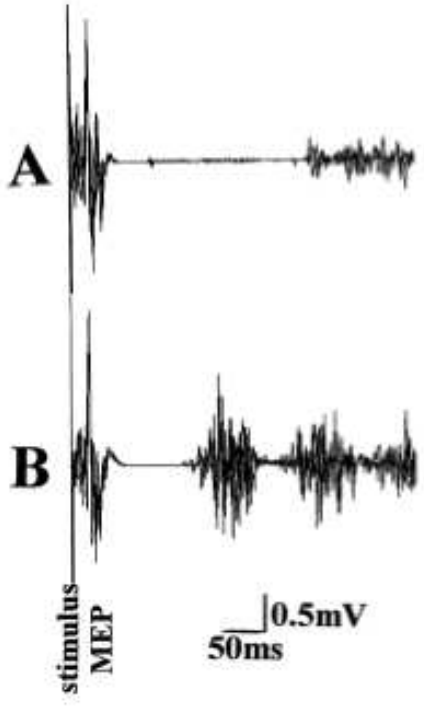

FIGURE 2.3: Silent period produced by single transcranial magnetic stimulation (TMS) in one patient. (A) TMS in the affected hemisphere; (B) TMS in the unaffected hemispheres. In each trace, 5 recordings are superimposed.

The CSP is also prolonged in patients when the affected hemisphere is stimulated, showing that cortical disinhibition on the unaffected side may occur simultaneously with an enhanced cortical inhibition on the affected side (See Fig.2.3).

Importantly for the purpose of this thesis, the CSP can be used as a prognostic parameter after stroke, and as a means to gauge the effectiveness of rehabilitation. The duration of the silent period in a normal individual depends mainly on the stimulation intensity, while the level of background contraction is of less importance [57]

Most reports seem to agree that the threshold for evoking a Silent Period is relatively stable in patients, especially if expressed as relative to motor threshold intensities. However, the duration of the Silent Period is prolonged, at least in the acute phase os the affected site. Traversa *et al.* ([68]), showed that the Silent Period shortens during recovery and there is some suggestion that the amount of such shortening correlates with the recovery of hand function[20]. More recent studies support this correlation as a good prognostic indicator and, thus, as a good indicator of the capabilities of the therapy [27, 28]. However, this shortening of the Silent Period is strongly dependent on the type of therapy chosen.

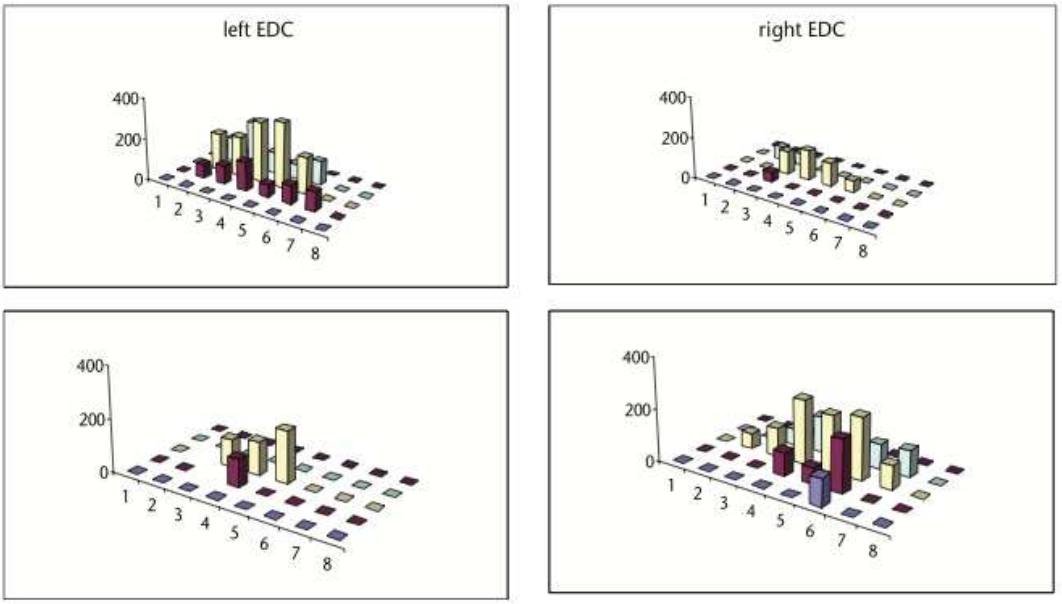

FIGURE 2.4: Three-dimensional map of the representation of the extensor digitorum communis (EDC) muscle after stimulation of a patient who, 6 months previously, had an infarct on the left internal capsule. The map was made before (top row) and after (bottom row) 3 weeks of constraining-induced movement therapy. The x-y grid represents the surface of contralateral scalp, divided into 1-cm squares. The height of each bar indicates the mean EMG response (in microvolts).

Chapter 3

Transcranial Magnetic Stimulation

3.1 Principles of Transcranial Magnetic Stimulation

3.1.1 Introduction

Over the last decade, Transcranial Magnetic Stimulation (TMS) has become one of the most commonly used neuroscience techniques to study the workings of the brain. Differently to the other classical neuroimaging techniques, such that EEG, fMRI, PET, MEG and so on, TMS allows neuroscientists to directly stimulate the brain safely without producing any discomfort to the patient.

The study of the circuitry in the brain has come a long way since the 19^{th} century, when Santiago Ramon y Cajal discovered how the cells in the nervous system were connected, providing detailed descriptions of cell types associated with neural structures and producing excellent depictions of neural structures and their connectivity (See Figure 3.1). This so called neuron doctrine stems from the fundamental idea that the nervous system is made of discrete individual cells called neurons [62]. These neurons communicate with one another by means of long protoplasmic fibers called axons, which carry trains of signal pulses called action potentials to distant parts of the brain or body and target them to specific recipient cells.

The brain as a whole can be extremely complex. The cerebral cortex, the most external layer of the human brain, contains roughly around 15-33 billions of neurons [56], linked by up to 10.000 synaptic connections each. The communication between neurons is conveyed through chemical signals by so-called neurotransmitters. Depending on which neurotransmitter the neuron is specialized, neurons can be excitatory or inhibitory. These features make the brain a device guided by nonlinear relations, in computational terms. For example, noradrenaline and acetilcoline are neurotransmitters of excitatory neurons, whereas gamma-aminobutiric acid is a neurotransmitter of inhibitory neurons.

Transcranial magnetic stimulation provides a way to change the excitability at circuitry level. For example, if a population of neurons is not activated, a magnetic pulse generates an electrical field capable of activating the default state neurons. Otherwise, if a populations of neurons presents sustained activity, then the magnetic pulse will cause a disruption on such activity by activation of inhibitory circuits [33]. So then, TMS can be used to study and explain the complexity of natural neural networks.

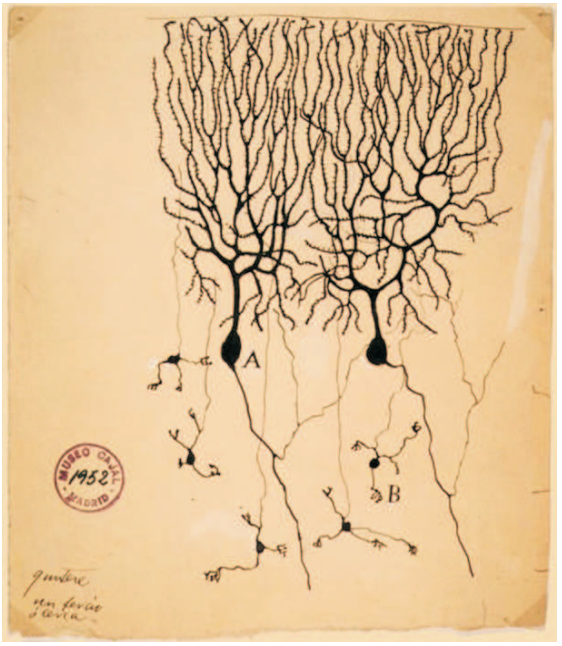

FIGURE 3.1: Drawing of Purkinje cells (A) and granule cells (B) from pigeon cerebellum by Santiago Ramon y Cajal, 1899. Instituto Santiago Ramon y Cajal, Madrid, Spain.

This chapter is divided into four sections. The first one aims to provide some historical background of brain stimulation. The second section describes the theory and technical issues of TMS, focusing on physical principles and devices. The third one outlines one of the most widely studied applications of TMS, which is motor cortex stimulation. Finally, the fourth section concludes with a brief introduction about the cortical silent period, describing the physiology of this motor inhibition parameter, of central interest to the current thesis.

3.1.2 Historical Introduction to Brain Stimulation

In the first half of the 19^th century, it was generally believed that there was no localization of function within the cerebral cortex. During the 1860s this view was challenged by the clinical observations of Brocca on patients with left frontal lesions and disturbances of language, and those of Hughlings Jackson on patients with focal seizures. These observations pointed to the possibility that different activities were localized in different parts of the brain. In particular, Jackson's observation on what are now known as Jacksonian seizures, led him to conclude that event within motor areas of the brain, different parts of the body were mapped into different anatomical areas. In this way, he could explain the spread of a seizure which began in the fingers and the progressed to the upper arm, and shoulder and then to the face and trunk.

These speculations were later confirmed by the experiments of Fritsch and Hitzig ([29]), and by Ferrier ([26]). Fritsch and Hitzig applied galvanic (direct current) stimulation to the cortex through a pair of electrodes placed 2-3 mm apart at an intensity sufficient to evoke a sensation when applied to the experimenter's tongue. This type of stimulation produced brief muscle twitches on the contralateral side of the body. The area from which the twitches could be evoked was a relatively circumscribed zone in the frontal part of the dog's brain. Ferrier also confirmed the localization of the motor areas of the brain, and also noted that different parts of the body were represented in different areas of the motor strip. He used galvanic and faradic (alternating

current stimulation at 30-40 Hz) stimulation, and commented that the latter produced more prolonged movement than the brief twitches elicited by galvanic stimuli.

In 1874, Bathlolow delivered the first description of the electrical excitability of the human cortex [6] . The observation that motor responses could be elicited by direct stimulation of the brain of man was subsequently reconfirmed by several neurosurgeons around the turn of the century. Despite these advances made by neurosurgeons, no method of brain stimulation through the intact skull in normal, behaving subjects, was until recently available. Several attempts were made in the 1950s and most were unsuccessful. Only one report claimed to have achieved transcraneal stimulation in man Gualtierotti and Paterson ([32]) used bipolar stimulation with trains of electrical pulses lasting for up to 40s, delivered through electrodes held in place by rubber bands. They reported movements of the contralateral hand and arm which were tonic in nature and persisted throughout the period of stimulation. Because this form of repeated stimulation was rather painful it appears that ether or nitrous oxide anaesthesia was used in their subjects.

In 1980, Merton and Morton [50] succeeded in producing movement of contralateral limb muscles by electrical stimulation of the motor cortex through the scalp. They discovered that the secret of the technique was to use a large single shock, rather than small repetitive shocks as those used conventionally on exposed cortex. Finally, in 1985, Barker, Jalinous and Freeston ([4]), stablished that a magnetic field also was capable of activating the human motor cortex through the skull. This method, virtually painless, has since achieved widespread acceptance.

3.1.3 Technical and physiological aspects of Transcranial Magnetic Stimulation

The bases of TMS are related at large to the theory of electro-magnetic induction, and are generally described by Maxwell's equations. Because the electromagnetic fields associated with TMS are of low frequency, the quasi-static approximation of the equations can be applied to the computation of the tissue-induced fields and currents.

As described in [61] A time-varying current pulse, produced using a very wired coil, produces a magnetic field according to the Biot-Savart law. The time varying magnetic field, in turn, induces an electric field according to Faraday's law. The induced electric field moves charges in the direction of its field lines (Fig.3.2). If the coil is parallel to the surface of the head, no surface charges appear due to the induction, and the computation of the electric field inside the conductor is simple. Otherwise, charges accumulate at the conductor surface as well as the interfaces between tissues with different conductivity, generating a secondary electric field. The intensity of the magnetic field can be represented by flux lines around the coin and is measured in tesla (T). The orientation of the magnetic field is perpendicular to the coil and, for currently available devices, can reach values of up to 4 T even if average values of maximal stimulating intensities around 2 T are currently provided by commercially available devices.

The physical expression leading the basic principle of TMS is as follows:

$$E = \frac{\delta A}{\delta \tau} - \Delta V$$

The expression for the total induced electric field E inside a conductor has a term due to induction, represented by vector potential A, and a term from surface changes, represented by scalar potential V.

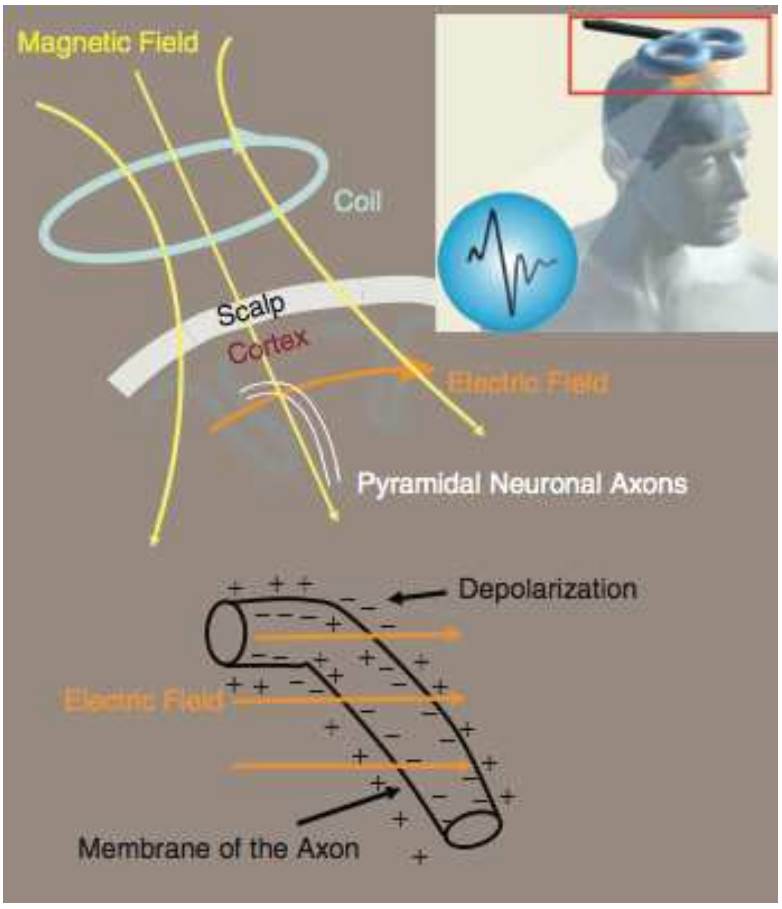

FIGURE 3.2: The electric current flowing in the TMS coil induces a magnetic field, with penetrates the layers of tissues separating the brain from the outside (scalp). The magnetic field induces an orthogonal electric field in the underneath brain cortex, which causes the movement of ions. Neuron axons orthogonal to the electric field are maximally perturbed in their membrane potential and undergo axon depolarization or hyperpolarization depending on axon orientation with respect to the current flow direction

Because the total induced electric field is strongest at the boundaries of any homogenous conductor compartment, the stimulation effect of TMS in the brain is concentrated at the cortical surface. The electric field induced in the cortical tissue causes the cell membranes to either depolarize or hyperpolarize. If the depolarization of the membrane overcomes its threshold, an action potential is generated. For example, one classical assumption is that the activation of pyramidal neurons by TMS occurs predominantly via interneurons in superficial cortical layers. Macroscopically, the locus of activation in the brain seems to be where the induced field is maximal. Focal activation is achieved by using a figure-of-eight coin or a double-cone coil with two loops, in which the current flows in opposite directions. The induced electric field peaks at the intersection the the coil windings.

The stimulation field experienced by a neuron is of a duration equal to the first phase of the time-varying waveform of the magnetic field. A greater amount of stored energy is required for longer-duration stimuli to achieve the same change in transmembrane voltage. Therefore, short pulses with rise times of less than $100\mu s$ are usually applied.

Large magnetic field pulses need to be generated in order to induce intra cortical electric fields of enough amplitude and duration to provoke the concurrent activation of the surrounding neural

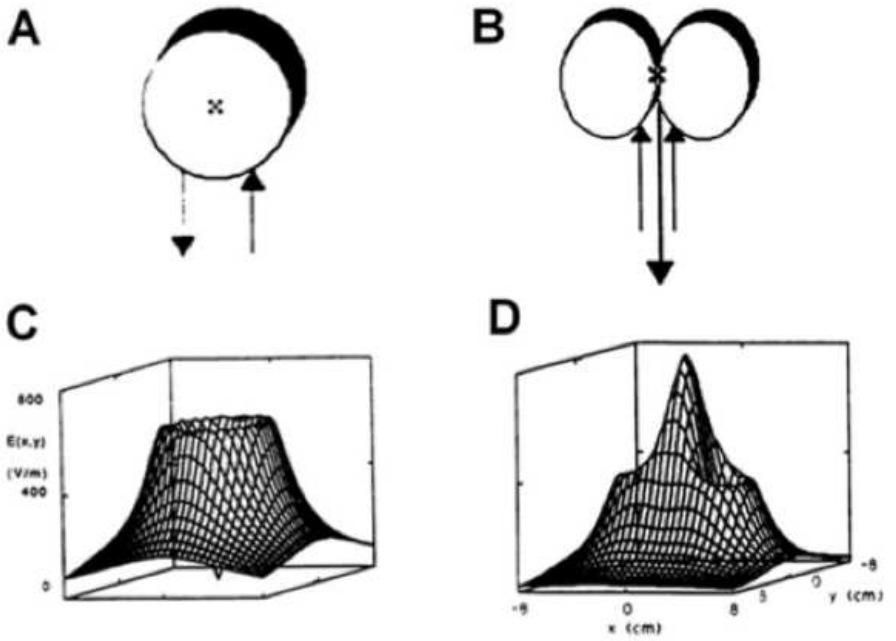

FIGURE 3.3: (a) Circular coil with the produced stimulating field. Notice that the part of maximal intensity is localized along the circumference (c), while the one with minimal intensity is located right in the center of the coil. (b) Figure-of-eight coil. Notice that the intensity of maximal stimulation (d)corresponds to the point of conjunction between the two wings.

tissue. To release sufficient energy in a very short time, a magnetic stimulator works by charging one or more energy storage capacitors (Fig.3.4) and then rapidly transferring this stored energy from the capacitor(s) to the stimulating coil and it discharges. In this way, a very high electrical current with a peak value of more that $5.000A$ flows from the capacitor(s) through the stimulation coil generating the required magnetic fiend.

The difficulty in producing magnetic nerve stimulators is related to the high discharge currents, voltages and power levels involved in producing the brief magnetic pulse. Typically, $500J$ of energy has to be transferred from the energy storage capacitor into the stimulation coil in around $100\mu s$. As power, measured in watts, is equivalent to joules per second, the power output of a typical magnetic stimulator during the discharge phase is $5Mw$ $(5.000.000MV)$. As a curiosity, such electrical power could provide the electricity used by $1,000$ houses over $1ms$.

During the discharge, the energy that was initially stored in the capacitor in the form of electrostatic charge is suddenly converted into magnetic energy in the stimulating coil in approximately $100\mu s$. This rapid rate of energy transfer produces a time-varying magnetic field buildup that induces tissue currents in the proximity of the coil in the order of 1-20 mA/cm^2. However, the amount of thermal energy delivered to the surrounding tissues due to magnetic stimulation is very small. At maximal output, assuming a maximal stimulus repetition rate of one pulse every $3s$ the average power deposited in the brain is calculated to be less than $2mW$.

The stimulating coil, normally housed in a molded plastic package, consist of one or more tightly wound and well-insulated copper coils together with other electronic circuitry such as temperature sensors and safety switches. At the present time, most commercial magnetic stimulator are supplied with a circular coil of 5-10 cm diameter. Different coil types are nowadays

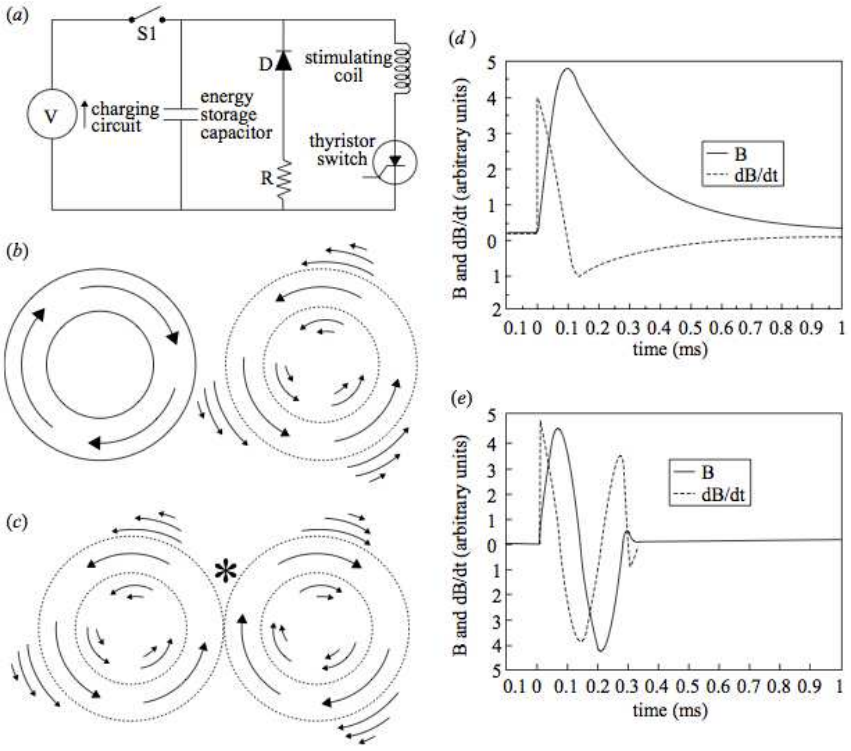

FIGURE 3.4: (a) Diagram of the circuit of a simple magnetic stimulator. (b) Outline of a simple round stimulating coil. The arrows indicate the direction of current flow in the coil. To its right is shown a schematic of the induced electric field in the tissue directly beneath the coil. The induced current is opposite so that in the stimulating coil. (c) Schematic of the induced current in the brain beneath a figure of eight coil. The maximum current is in the vicinity of the virtual cathode, shown by the asterisk. (d) The time-course of the magnetic field and induced electric field waveforms (solid and dotted lines, respectively) beneath the centre of a simple round coil stimulation driven by standard rate Magstim 200 machine. (e) The same when driven by a Magstim Rapid stimulator.

available, each with their own advantages and disadvantages. Large coils cannot produce very focal stimulation but can penetrate relatively deep in the brain.

Although the circular coil is very useful as a general coil, it does not provide a very defined site of stimulation. For example, with a standard round coil, the induced current in the brain flows in an annulus, underneath the coil, which is usually some $8 - 12cm$ in diameter. Clearly, a large volume of neural tissue may be activated by such a device. Increases in the focality of stimulation can be achieved by tilting these circular coils so that their plane is not tangential to the skull at that point. The greater the angle between the skull and the coil, the more focal is the stimulation.

Figure-of-eight coils, also called *butterfly* or double coil, induce an electric field under the junction region of the eight, which is twice as large as that under the two wings. These coils can sustain larger currents because of the lower induction with respect to circular coils with the same number of rings (see Fig.3.3).

3.1.4 Transcranial Magnetic Stimulation of the Motor Cortex

Transcranial magnetic stimulation can be applied in several ways and at different levels of the nervous system: in the spinal cord, brainstem and peripheral nerves, and, of course, at cortical level [65]. However, TMS of the cerebral cortex is quite different from that of other parts of the nervous system. Stimulation of the peripheral nerves, spinal roots, and the afferent and efferent tracts at the cervicomedullary junction, essentially affects nerve axons. In contrast, the TMS of the motor cortex can evoke D waves registered in a electromyography device (EMG), representing direct stimulation of the cortico spinal axon, as well as I waves that arise from trans-synaptic activation of corticospinal neurons. This complex is caller Motor Evoked Potential (MEP), an example of which can be seen in Fig.3.5.

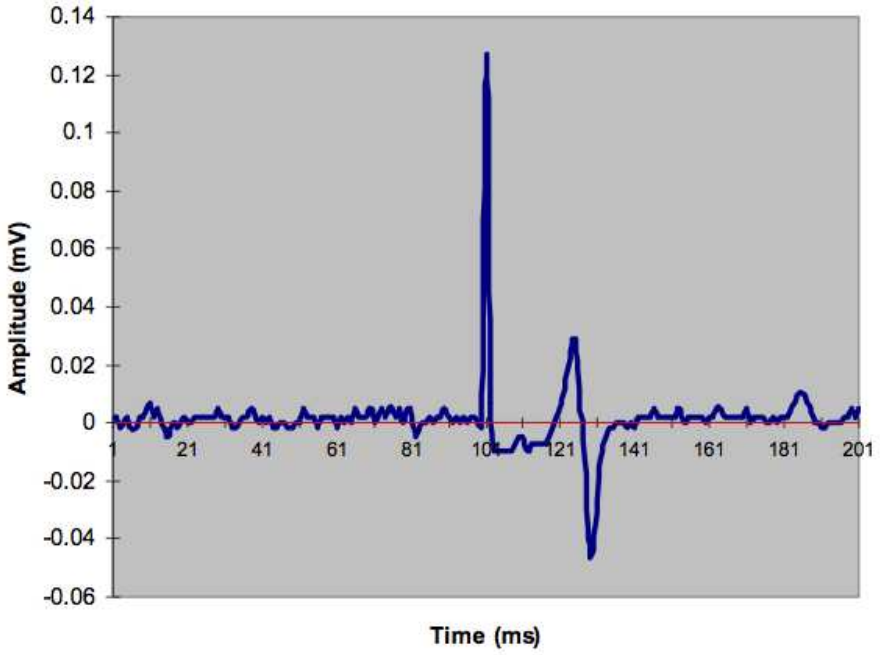

FIGURE 3.5: A representative motor-evoked potential (MEP) amplitude beginning approximately 20 ms after the stimulus artifact (huge amplitude)

Due to this intervening synapse between descending fibers and spinal motoneurones, and the possibility of evoking more than one descending volley within the pyramidal tract, the EMG responses to cortical stimulation are more complex than those following peripheral stimulation. On this side, two important effects can be observed: The first one is the latency; one of the clearest differences between muscle responses to peripheral nerve and cortical stimulations is that cortical latencies are shorter by as much as 4 ms in active compared with relaxed muscles. Since responses in active muscles are always larger than those in relaxed muscles for a given intensity of stimulation, one possible reason for the latency difference could be the size of of motor unit recruitment. The other important effect concerns the size and complexity of surface EMG responses. At relatively high levels, cortical stimulation produces EMG responses which are generally quite simple and comparable with those following stimulation of peripheral nerves [61].

3.1.5 Inhibitory effects and the Cortical Silent Period

With cortical stimulation, not only excitatory effects but also inhibitory effects can be elicited. This characteristic is used to investigate cerebral functions other than those of the motor cortex. [33].

When an individual is instructed to maintain muscle contraction and a single suprathreshold TMS pulse is applied to the motor cortex contralateral to the target muscle, the electromyographic activity is arrested for a few hundred milliseconds after the MEP [39] (Fig.3.6).

This period of electromyographic suppression is referred to as a silent period, normally defined as the time from the end of the MEP to the return os voluntary electromyographic activity. However, it is difficult at times to define the end of the MEP, especially in patients with corticospinal tract dysfunction. In order to circumvent this difficulty, some investigators have defined the silent period as the interval from stimulus delivery to the return of voluntary activity [69]

Most of the silent period is believed to be due to inhibitory mechanisms such as Renshaw cells inhibitions, which are thought to contribute to the first 50-60 ms os this suppression [19]. The silent period is likely to be mediated by the neurotransmitter $GABA_B$ receptors. Supporting this hypothesis, silent periods of abnormally short or long duration are observed in patients with various movement disorders [9]. For instance, patients with amyotrophic lateral sclerosis often show a shortened duration of silent periods due to impairment of intracortical inhibition that can be reversed by antiglutamatergic drugs; these findings provide insights into the pathophysiology of this disease [17].

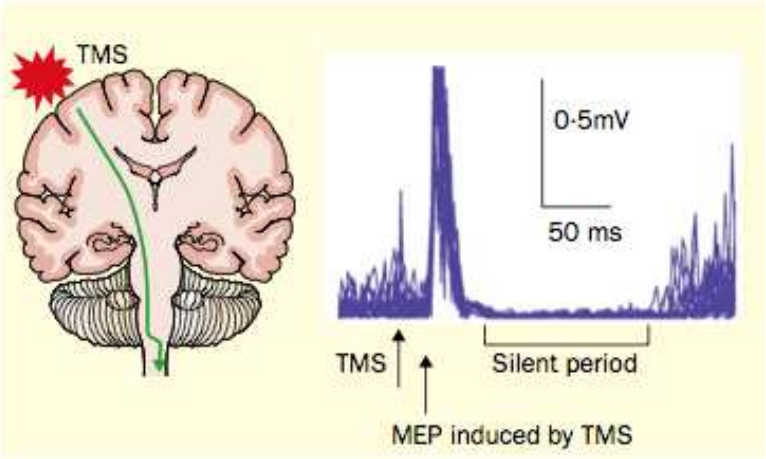

FIGURE 3.6: The effect of TMS on a motor area that represents a voluntarily contracted hand. Stimulation of the left motor cortex Elicits MEPs in the contralateral First Dorsal Interosseus, which is followed by a supression of tonic voluntary activity. Rectified 15 EMG responses are superimposed

Classen and co-workers [21] investigated patients after acute stroke, who showed hemiparesis and a long duration of the silent period, but normal MEP amplitude in the affected side. These patients had impaired movement initiation, inability to maintain a constant force, and impaired movement of individual fingers that resembled motor neglect. The silent period duration decreased with clinical improvement. This study suggests that, among patients with hemiparetic stroke, there is a subgroup whose motor disorders, involving features of motor neglect, are caused more by exaggerated inhibitory mechanisms in the motor cortex than by a direct corticospinal disorder.

Chapter 4

Machine Learning Background

4.1 Machine Learning Background

4.1.1 Introduction

Machines learn, humans decide. This is, in a nutshell, one of the conceptual drivers of this thesis. Humans seamlessly process information and perform pattern recognition. This is, they are capable of extracting meaningful knowledge from a complex mix of streaming raw sensory data, with little or none at all awareness of the mechanisms involved. This pattern recognition ability is in turn embedded into ongoing learning and inference systems.

Our research deals with a problem of pattern recognition from multivariate time series of electrophysiological data. Neurology experts and clinicians alike should have learnt over time to extract meaningful knowledge from this type of data. Nevertheless, the pattern recognition capabilities of the (trained) human brain are still limited by evolutionary constraints. In other words, human pattern recognition has evolved to match the requirements of the environment we live in, not to match the requirements of data analysis in general.

Machine learning techniques come handy to overcome such limitations. They are, in a general way, computational simulations of learning processes, in which learning may come in different flavors: it can be the result of a training process in which we seek to model the relation between a set of observed data and their corresponding outcome; or it can be the emergent result of an autonomous, self-organizing process. In one way or another, importantly, machine learning techniques can deal with large data sets of high-dimensional data, well beyond the capabilities of the human expert.

One of the central problems in machine learning and pattern recognition is that of finding low-dimensional representations of multivariate data residing on high-dimensional data spaces. Latent variable models address this problem by representing information from an observable, usually high dimensional data space, in an unobservable or latent, usually low-dimensional, space. In [64], these models are typified as belonging to different but overlapping categories: projection models, generative models, and other related models.

Projection models aim for the projection of data points residing in $\Re^D$, onto a hyperplane, $\Re^L$, with $\Re^L \subseteq \Re^D$, where $L \leq D$. The most common projection model (and, by far, the most

widely used) is Principal Component Analysis (PCA) [36], although other models, such as principal curves and surfaces [34], auto-associative feed-forward neural networks [42], and kernel based PCA [22], are also of common use.

Generative models are defined stochastically and try to estimate the distribution of data by defining a density model with low intrinsic dimensionality within the multivariate data space. Possibly, Factor Analysis (FA) [5, 44] is the most widely used generative model. It must be noted though, that FA is sometimes confused with rotated variations of PCA and both are used in similar applications.

Most of the interest in generative models stems from the fact that they fit naturally into the Statistical Machine Learning category and, in general, to the much wider framework of probability theory and statistics. Furthermore, generative models can directly make use of well-founded techniques for fitting them to data, combining different models, missing data imputation, outlier detection, etc.

Generative Topographic Mapping (GTM) is a non-linear generative model introduced in [12]. In short, it was defined to retain all the useful properties of Kohonen's Self-Organizing Maps (SOM) [41], such as the simultaneous clustering and visualization of multivariate data, while eluding most of its limitations through a fully probabilistic formulation. Its probabilistic setting has enabled the definition of principled extensions for hierarchical structures [67], missing data imputation [53], adaptive regularization [11, 73], discrete data modelling [13, 31], robust outlier detection and handling [74], and semi-supervised learning among others.

GTM-Through Time (GTM-TT) [10] is an extension of GTM suited to deal with the unsupervised analysis of multivariate time series. Given that it is defined as a constrained Hidden Markov Model (HMM) [59], GTM-TT explicitly accounts for the violation of the independent identically distributed (i.i.d) condition. This model has been assessed using artificial and real multivariate time series in [54].

In this chapter, we first describe the original GTM and its extension for the modeling of multivariate time series: GTM-TT. A Bayesian approach to the definition of GTM-TT is then outlined. This Bayesian formulation should deal effectively with the problem of overfitting. Finally, a recently developed Variational approach for Bayesian GTM-TT [55] is summarily introduced.

4.1.2 The Original GTM

The neural network-inspired GTM is a nonlinear latent variable model of the manifold learning family with sound foundations in probability theory. It performs simultaneous clustering and visualization of the observed data through a nonlinear and topology-preserving mapping from a visualization latent space in $\Re^L$ (with L being usually 1 or 2 for visualization purposes) onto a manifold embedded in the $\Re^D$ space, where the observed data reside. The mapping that generates the manifold is carried out through a generalized regression function, given by

$$\mathbf{y} = \mathbf{W}\mathbf{\Phi}\left(\mathbf{u}\right) \tag{4.1}$$

where $\mathbf{y} \in \Re^D$, $\mathbf{u} \in \Re^L$, $\mathbf{W}$ is the matrix that generates the mapping, and $\mathbf{\Phi}$ is a matrix with the images of S basis functions ϕ_s (defined as radially symmetric Gaussians in the original formulation of the model). To achieve computational tractability, the prior distribution of $\mathbf{u}$ in latent space is constrained to form a uniform discrete grid of K centres, analogous to the layout of the SOM units, in the form of a sum of delta functions:

$$p(\mathbf{u}) = \frac{1}{K}\sum_{k=1}^{K}\delta(\mathbf{u} - \mathbf{u}_k) \tag{4.2}$$

Defined in this way, GTM can also be understood as a constrained mixture of Gaussians. That is, a special case of a Gaussian mixture model that is adapted to provide high-dimensional data visualization. In each component of the mixture, is generated a density model in data space. Assuming that the observed data set $\mathbf{X}$ is constituted by N independent, identically distributed (i.i.d.) data points $\mathbf{x}_n$, leads to the definition of a complete likelihood in the form:

$$p(\mathbf{X}|\mathbf{W},\beta) = \left(\frac{\beta}{2\pi}\right)^{ND/2}\prod_{n=1}^{N}\left\{\frac{1}{K}\sum_{k=1}^{K}\exp\left(-\frac{\beta}{2}\|\mathbf{x}_n - \mathbf{y}_k\|^2\right)\right\} \tag{4.3}$$

where $\mathbf{y}_k = \mathbf{W}\mathbf{\Phi}(\mathbf{u}_k)$. From Eq. 4.3, the adaptive parameters of the model, which are $\mathbf{W}$ and the common inverse variance of the Gaussian components, β, are usually optimized by Maximum Likelihood (ML) using the Expectation-Maximization (EM) algorithm [7]. Details of this calculations can be found in [12].

As mentioned, the GTM is embodied with clustering and visualization capabilities that are akin those of the SOM. Data points can be summarily visualized in the low-dimensional latent space(1 or 2 dimensions) of GTM by means of the *posterior-mode* projection [11], defined as

$$k_n^{mode} = \arg\max_{\{k_n\}} r_{kn}, \tag{4.4}$$

which also provides an assignment of each data point $\mathbf{x}_n$ to a cluster representative $\mathbf{u}_k$. The distribution of the *responsibility* over the latent space of states can also be directly visualized. Another possibility of visualization is, for each data point $\mathbf{x}_n$, to plot the mean of the posterior distribution in latent space,

$$\mathbf{u}_n^{mean} = \sum_{k=1}^{K} r_{kn}\mathbf{u}_k, \tag{4.5}$$

known as *posterior-mean* projection.

4.1.3 The GTM Through Time Method

The data mining of multivariate time series has long ago become an established research area ([18]). Methods dealing with this problem have stemmed from both traditional statistics and machine learning field, using neural networks as a fruitful approaches ([75]).

These methods usually consider the problem as supervised, being prediction the main goal of the analysis. In comparison, little research has been devoted to methods of unsupervised clustering for the exploration of the dynamics of multivariate time series.

Some of the most interesting time series clustering results have been obtained with different variants of SOM models in diverse contexts although, in general, without accounting for the violation of the i.i.d condition.

The GMT Through Time (GTM-TT: [10]) is an extension of GTM suited to deal with the unsupervised analysis of multivariate time series. It explicitly accounts for the violation of the i.i.d. condition, given that it is defined as a constrained HMM.

GTM-TT can be considered as a GTM model in which the latent states are linked by transition probabilites, in a similar fashion to Hidden Markov Models. In fact, GTM-TT can be understood as a topology constrained HMM.

Assuming a sequence of N hidden states $\mathbf{Z} = \{z_1, z_2, \ldots, z_n, \ldots, z_N\}$ and the observed multivariate time series $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n, \ldots, \mathbf{x}_N\}$, the probability of the observations is given by:

$$p(\mathbf{X}) = \sum_{\text{all } \mathbf{Z}} p(\mathbf{Z}, \mathbf{X}) \tag{4.6}$$

where $p(\mathbf{Z}, \mathbf{X})$ defines the complete-data likelihood as in HMM models [59] and takes the following form:

$$p(\mathbf{Z}, \mathbf{X}) = p(z_1) \prod_{n=2}^{N} p(z_n | z_{n-1}) \prod_{n=1}^{N} p(\mathbf{x}_n | z_n) \tag{4.7}$$

The model parameters are $\mathbf{\Theta} = (\boldsymbol{\pi}, \mathbf{A}, \mathbf{Y}, \beta)$ where $\boldsymbol{\pi} = \{\pi_j\} : \pi_j = p(z_1 = j)$ are the initial state probabilities, $\mathbf{A} = \{a_{ij}\} : a_{ij} = p(z_n = j | z_{n-1} = i)$ are the transition state probabilities, and

$$\{\mathbf{Y}, \beta\} : p(\mathbf{x}_n | z_n = j) = \left(\frac{\beta}{2\pi}\right)^{D/2} \exp\left(-\frac{\beta}{2} \|\mathbf{x}_n - \mathbf{y}_j\|^2\right)$$

are the emission probabilities, which are controlled by spherical Gaussian distributions with common inverse variance β and a matrix $\mathbf{Y}$ of K centroids $\mathbf{y}_j$, $1 \leq j \leq K$.

For mathematical convenience, it is useful defining a state in the vectorial form $\mathbf{z}_{j,n}$ such that it returns 1 if z_n is in state j, and zero otherwise. Using this notation, the initial state probabilities, the transition state probabilities and the emission probabilities are defined as:

$$p(z_1 | \pi) = \prod_{j=1}^{K} \pi_j^{\mathbf{z}_{j,1}} \tag{4.8}$$

$$p(z_n | z_{n-1}, \mathbf{A}) = \prod_{i=1}^{K} \prod_{j=1}^{K} a_{ij}^{\mathbf{z}_{j,n} \mathbf{z}_{i,n-1}} \tag{4.9}$$

$$p(\mathbf{x}_n | z_n, \mathbf{Y}, \beta) = \left(\frac{\beta}{2\pi}\right)^{D/2} \prod_{j=1}^{K} \left\{\exp\left(-\frac{\beta}{2} \|\mathbf{x}_n - \mathbf{y}_j\|^2\right)\right\}^{\mathbf{z}_{j,n}} \tag{4.10}$$

Eqs. 4.8 to 4.10 lead to the definition of the complete data log-likelihood as:

$$\begin{aligned}\ln p\left(\mathbf{Z},\mathbf{X}|\mathbf{\Theta}\right) &= \sum_{j=1}^{K}\mathbf{z}_{j,1}\ln\pi_j \\ &+ \sum_{n=2}^{N}\sum_{i=1}^{K}\sum_{j=1}^{K}\mathbf{z}_{i,n-1}\mathbf{z}_{j,n}\ln a_{ij} \\ &+ \frac{ND}{2}\ln\left(\frac{\beta}{2\pi}\right) \\ &- \frac{\beta}{2}\sum_{n=1}^{N}\sum_{j=1}^{K}\mathbf{z}_{j,n}\left\|\mathbf{x}_n-\mathbf{y}_j\right\|^2 \end{aligned} \tag{4.11}$$

Parameter estimation can be accomplished in GTM-TT by maximum likelihood using the EM algorithm, in a similiar fashion to HMMs. Details can be found in [10].

4.1.4 Bayesian GTM Through Time

Although the ML framework is widely used for parameter optimization, it shows two important weakness:

- Its maximization process does not take into account the model complexity
- It tends to overfit the model to the training data

The complexity in GTM-TT is related with several parameters, such that the number of hidden states, their degree of connectivity and the dimension of the hidden space. For visualization processes, the dimensions of the hidden space is limited to be three or less. The number of hidden states and the maximum number of possible state transitions are strictly correlated by a squared power. In order to solve the overfitting problem, the complexity has been limited by restricting the number of possible state transitions [10] or by fixing the transition state probabilities a priori [38]. The alternative technique of cross-validation is very expensive computationally speaking and it requires large amounts of data to obtain low-variance stimates of the expected test errors. To control overfitting and model complexity a full Bayesian reformulation of GTM-TT was recently proposed: The Variational Bayesian GTM-TT (henceforth VBGTM-TT [52, 55].

The Bayesian approach treats the parameters as unknown quantities and provides probability distributions for their priors. Bayes' theorem can be used to infer the posterior distributions over the parameters. The model parameters can thus be considered as hidden variables and integrated out to describe the marginal likelihood as:

$$\begin{aligned} p\left(\mathbf{X}\right) &= \int p\left(\mathbf{\Theta}\right)p\left(\mathbf{X}|\mathbf{\Theta}\right)d\mathbf{\Theta}, \\ &\text{where } \mathbf{\Theta} = \left(\boldsymbol{\pi},\mathbf{A},\mathbf{Y},\beta\right) \end{aligned} \tag{4.12}$$

If an independent distribution is assumed for each parameter, then:

$$p(\boldsymbol{\Theta}) = p(\boldsymbol{\pi})\, p(\mathbf{A})\, p(\mathbf{Y})\, p(\beta) \tag{4.13}$$

Taking into account Eqs. 4.6, 4.12 and 4.13, the marginal likelihood in GTM-TT can be expressed, similarly to HMM [8], as:

$$\begin{aligned} p(\mathbf{X}) = \int p(\boldsymbol{\pi}) \int p(\mathbf{A}) \int p(\mathbf{Y}) \int p(\beta) \\ \sum_{\text{all } \mathbf{Z}} p(\mathbf{Z}, \mathbf{X} | \boldsymbol{\pi}, \mathbf{A}, \mathbf{Y}, \beta)\, d\beta d\mathbf{Y} d\mathbf{A} d\boldsymbol{\pi} \end{aligned} \tag{4.14}$$

Although there are many possible prior distributions to choose from, the conjugates of the distributions defined in Eqs. 4.8 to 4.10 are a reasonable choice. In this way, a set of prior distributions is defined as follows:

$$\begin{aligned} p(\boldsymbol{\pi}) &= \mathtt{Dir}\left(\{\pi_1, \ldots, \pi_K\} | \boldsymbol{\nu}\right) \\ p(\mathbf{A}) &= \prod_{j=1}^{K} \mathtt{Dir}\left(\{a_{j1}, \ldots, a_{jK}\} | \boldsymbol{\lambda}\right) \\ p(\mathbf{Y}) &= \left[(2\pi)^K |\mathbf{C}|\right]^{-D/2} \prod_{d=1}^{D} \exp\left(-\frac{1}{2} \mathbf{y}_{(d)}^T \mathbf{C}^{-1} \mathbf{y}_{(d)}\right) \\ p(\beta) &= \Gamma\left(\beta | d_\beta, s_\beta\right) \end{aligned}$$

where $\mathtt{Dir}(\cdot)$ represents the Dirichlet distribution; and $\Gamma(\cdot)$ is the Gamma distribution. The vector ν, the matrix λ and the scalars d_β and s_β correspond to the hyperparameters of the model which are fixed a priori. The prior over the parameter $\mathbf{Y}$ defines the mapping from the hidden states to the data space as a GP, where $\mathbf{y}_{(d)}$ is each of the row vectors (centroids) of the matrix $\mathbf{Y}$ and $\mathbf{C}$ is a matrix where each element is a covariance function that can be defined as

$$C(\mathbf{u}_i, \mathbf{u}_j) = \nu \exp\left(-\frac{\|\mathbf{u}_i - \mathbf{u}_j\|^2}{2\alpha^2}\right), \quad i, j = 1 \ldots K$$

The α parameter controls the flexibility of the mapping from the latent space to the data space. The vector $\mathbf{u}_j,\ j = 1 \ldots K$ corresponds to the state j in a latent space of usually lower dimension than that of the data space. Thus, a topography over the states is defined by the GP as in the standard GTM.

A graphical model representation of the proposed Bayesian GTM-TT can be seen in Fig. 1.

Unfortunately, Eq. 4.14 is analytically intractable. In the following section, we provide the details of its approximation using Variational inference techniques.

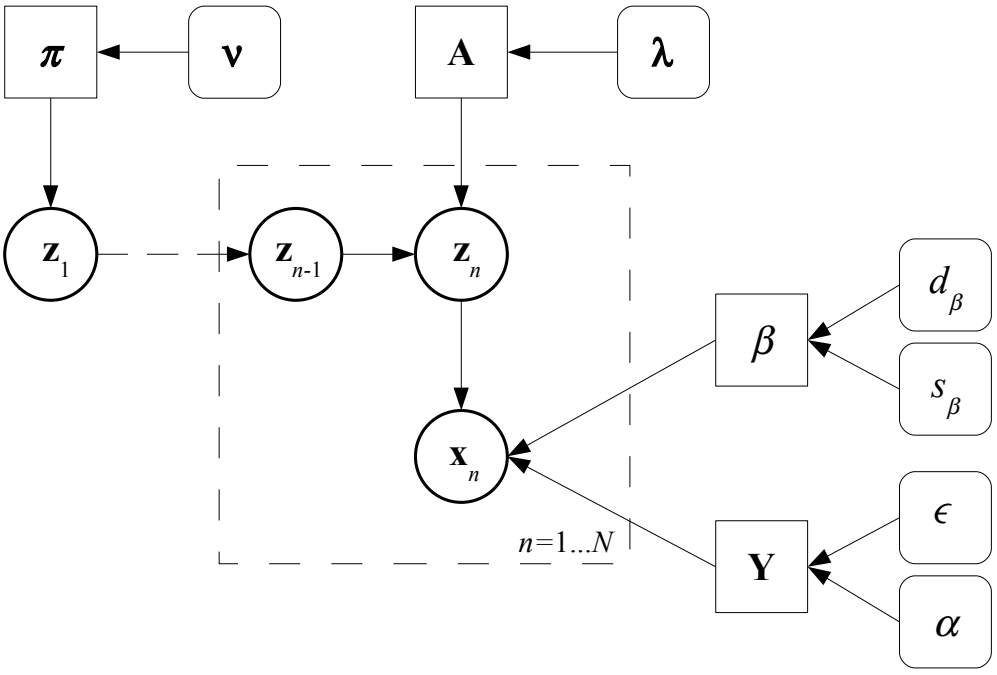

FIGURE 4.1: Graphical model representation of the Bayesian GTM-TT. Variables are noted by circles, paramenters, by squares and hyperparamenters, by rounded squares

4.1.5 A Variational Approach for Bayesian GTM-TT

Variational inference allows approximating the marginal log-likelihood through Jensen's inequality as follows:

$$\begin{aligned} \ln p(\mathbf{X}) &= \ln \int \sum_{\text{all } \mathbf{Z}} p(\mathbf{Z}, \mathbf{X}|\mathbf{\Theta})\, p(\mathbf{\Theta})\, d\mathbf{\Theta} \\ &\geq \int \sum_{\text{all } \mathbf{Z}} q(\mathbf{\Theta}, \mathbf{Z}) \ln \frac{p(\mathbf{Z}, \mathbf{X}|\mathbf{\Theta})\, p(\mathbf{\Theta})}{q(\mathbf{\Theta}, \mathbf{Z})} d\mathbf{\Theta} \\ &= F(q(\mathbf{\Theta}, \mathbf{Z})) \end{aligned}$$

The function $F(q(\mathbf{\Theta}, \mathbf{Z}))$ is a lower bound such that its convergence guarantees the convergence of the marginal likelihood. The goal in variational inference is choosing a suitable form for the approximate density $q(\mathbf{\Theta}, \mathbf{Z})$ in such a way that $F(q)$ can be readily evaluated and yet which is sufficiently flexible that the bound is reasonably tight. A reasonable approximation for $q(\mathbf{\Theta}, \mathbf{Z})$ is based on the assumption that the hidden states $\mathbf{Z}$ and the parameters $\mathbf{\Theta}$ are independently distributed, i.e. $q(\mathbf{\Theta}, \mathbf{Z}) = q(\mathbf{\Theta})\, q(\mathbf{Z})$. Thereby, a Variational EM algorithm can be derived [8]:

VBE-Step:

$$q(\mathbf{Z})^{(\text{new})} \leftarrow \underset{q(\mathbf{Z})}{\text{argmax}}\, F\left(q(\mathbf{Z})^{(\text{old})}, q(\mathbf{\Theta})\right) \tag{4.15}$$

VBM-Step:

$$q\left(\mathbf{\Theta}\right)^{(\text{new})} \leftarrow \underset{q(\mathbf{\Theta})}{\text{argmax}}\, F\left(q\left(\mathbf{Z}\right)^{(\text{new})}, q\left(\mathbf{\Theta}\right)\right) \tag{4.16}$$

The model is implemented by these two equations. Details of such implementation can be found in [52].

Chapter 5

Experiments

5.1 A novel approach to the estimation of the Contralateral Silent Period duration

5.1.1 Introduction

As described in Chapter 3, TMS of the contralateral motor cortex during voluntary muscle activity produces a motor evoked potential (MEP), followed by a period of cessation of EMG activity, known as the cortical silent period (CSP).

The early part of this CSP is at least partly related to decreased spinal motor neuron excitability, while the late part (about 50 ms) is due to intracortical inhibitory mechanisms [19, 69].

The CSP has been shown to be abnormal in many neurological and psychiatric disorders, reflecting altered cortical inhibition. It has thus arguably become a key to the understanding of the pathophysiology of these disorders.

How to demarcate both the start and end of the CSP is an open subject of scientific debate. The practice of fixing the onset of the CSP at the stimulus onset is generally well established. Instead, the determination of the end of the CSP is a very difficult task. The return of voluntary EMG activity as a marker of the end of the CSP is frequently imprecise.

Terms such as the absolute and relative CSP have been introduced to provide some uniformity when deriving this measure. The end of the absolute CSP is formally defined as the point in which any EMG activity returns. In contrast, the end of the relative CSP is defined as the time when the EMG activity approaches the pre-stimulus state [66]. Following either criterion, extraneous signals originating from a variety of sources such as movement artifacts can often cause deflections in the waveform that may be inaccurately interpreted as a return of voluntary motor activity, even after averaging many trials over the same subject.

As a result, the determination of the CSP can be subject to interpretative arbitrariness and often leads to discrepancies between raters. Given that CSP determination lacks the necessary objectivity that is a conventional requirement for any neurophysiologic measure, any technique capable to reliably automate this estimation procedure would be extremely beneficial, as it would both simplify and standardize the estimation procedure. If applied to the data acquired from patients undergoing motor rehabilitation from stroke, as in this thesis, such technique might help to discern with minimum ambiguity the rehabilitation progress.

Nilsson *et al* ([19]) endeavored to measure the CSP duration through a computer-automated approach. The end of the CSP was defined as the first epoch in which no significant difference, as compared to baseline conditions, was to be found. However, there are several drawbacks in this method: The main one is that it fails to account for the magnitude of such difference. Hence, a given epoch (2-4 ms of signal), may not necessarily differ statistically from baseline conditions yet still represent a period of partial cessation of voluntary motor activity and, therefore, be part of the CSP.

In another approach to this problem, Garvey *et al*. ([30]) reported an alternative mathematical approach for determining the CSP duration based on the mean consecutive difference (MCD) of the data points in the pre-stimulus EMG. The offset of the CSP was defined as the first data point of a 5 ms window in which at least 50% of the values are higher than the expression (mean-2.66$\times$MCD) for the pre-stimulus EMG. Using this approach, they found that the obtained CSP values closely approximated those obtained using the manual method. Again, there are several problems associated with this approach: considerable differences were found when measured in children (approximately 10 ms). This is in part due to the fact that the CSP in children is often short or directly non-existent. Also, the method is extremely sensible to muscle-artifacts and other sources of noise, making signal filtering necessary. This filtering affects the calculation of the pre-stimulus activation that is necessary to determine the offset of CSP.

5.1.2 The estimation of the CSP duration

5.1.2.1 The current standard for the estimation of the CSP duration

The current standard for the automated calculation of the offset of CSP was developed by Daskalakis *et al* ([23]). In this automated approach (hereafter referred to as Daskalakis' Method, or DM for short), a combination of filtering, squaring and threshold detection is used to calculate the CSP duration. The method has the advantage of being objective, as it avoids the potential biases and interpretation discrepancies inherent to the previously described, more conventional and visually-guided CSP measuring techniques. It has also been widely accepted for its simplicity of implementation, which is accessible to people with little or no computational background. In this approach, the end of the CSP is marked by the first return of any voluntary EMG activity, which was defined as the first time point at which the background EMG was restored to 25% of pre-stimulus EMG amplitude. An example of application of this method can be seen in Fig. 5.1.

Despite the reliability of this method to automate the determination of the CSP at short or long durations, its main limitations reside, first, in the inability to cope with more complex CSP cases where additional CSPs follow bursts of EMG activity; and secondly, in its bad performance in the presence of noisy data, which are commonplace in pathological EMG recordings. While acquiring the EMG signal of activity from patients who are not able to maintain a voluntary muscle contraction, or who suffer some sickness related to muscle control, spontaneous EMG activity during the CSP is likely to occur. This activity is not relevant to the analysis, because it is not related with disinhibition of pyramidal cells. Daskalakis' method cannot distinguish these cases. Moreover, the 50 Hz high-pass filtering condition in this method sometimes hides relevant EMG signal information related to the return of the voluntary contraction, and can result in an unreliable calculation of the CSP.

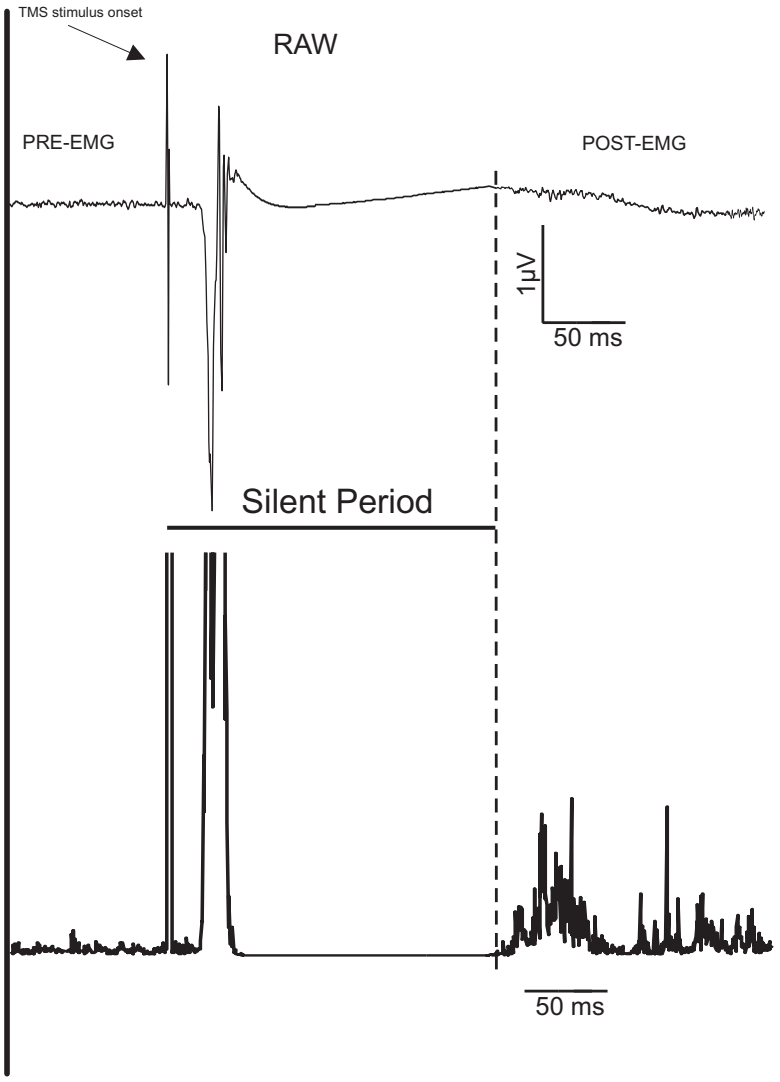

FIGURE 5.1: Transformation of raw data used in DM to detect and measure CSP. Stimulation levels are 150% of the resting motor threshold. The top trace in an example of raw recordings of a TMS-evoked MEP and CSP during tonic activation of the FDI muscle. After unprocessed traces are analyzed by DM, they become as shown in the bottom display. The calculation for the onset of the CSP is based on the stimulus onset. The threshold calculation for the offset of the CSP is based on the amplitude of the pre-stimulation EMG.

5.1.2.2 A novel approach: the index of variability

The GTM-TT model, described in the previous background chapter on Machine Learning methods, can facilitate the faithful identification and visualization of change-points and sudden transitions in MTS [55]. Change-points, in the low-dimensional visual data representation provided by the model, correspond to sudden jumps between often distant model states. Instead, subsequences of little variability over time will often clump in few model states or even remain in a single one over time.

In this thesis, we hypothesize that the CSP can be encapsulated by change points. In other words, we claim that the GTM-TT model should be able to unambiguously identify the CSP as a time subsequence of little variability, bounded by change points.

To test such hypothesis, and beyond the exploratory visualization of MTS that GTM-TT can provide, we need a well-defined measure of MTS variation that allows us to identify and quantify change-points. in GTM-TT, we expect sudden transitions to be accompanied by sudden increases of the model likelihood [**?**], so that the weighted mean of the emission probabilities of the model in logarithmic form could be a good candidate to consider as an Index of Variability (IV):

$$IV_n = -\sum_k r_{k,n} \ln p\left(\mathbf{x}_n | \mathbf{z}_n = k\right),$$

where $r_{k,n}$ is the responsibility (a posterior probability) taken by a hidden state $z_n = k$ out of K for each point x_n in the MTS.

Unfortunately, this measure can be affected by noise, and it will no reflect the advantages of the data regularization provided by the Variational Bayesian-GMT-TT. For this reason, a novel IV, namely the weighted-prototype IV ($wpIV$), is proposed for the latter model and defined as:

$$wpIV_n = \|\Omega_n^{mean} - \Omega_{n-1}^{mean}\| \ , \tag{5.1}$$

where $\|\cdot\|$ is the Euclidean distance and $\Omega_n^{mean} = \sum_{k=1}^{K} \langle z_{k,n} \rangle \mathbf{y}_k$. Here, the variational parameter $\langle z_{k,n} \rangle$ plays the same role as $r_{k,n}$ plays for the standard GTM-TT, and vector $\mathbf{y}_k$; $k = 1 \ldots K$ is the data prototype of state k in data space. Eq.5.1 is nothing but the weighted distance between the data prototypes representing two consecutive instants in the MTS, where each prototype can take at least partial responsibility for the representation of each instant of the MTS. The same distance measured between the observed data of the two consecutive instants would be of little use as any relevant information would be masked by noise.

By measuring the distance using the model-generated prototypes, we ensure that, provided the model manages to faithfully recover the underlying structure of the MTS (and VB-GTM-TT does this by avoiding overfitting while the standard GTM-TT cannot), the true change-points will be clearly detected.

5.1.3 The experimental setting

The previous sections should provide the boundaries for our experiments. Given that we are proposing a new technique -based on a novel Statistical Machine Learning model- that might be useful to estimate the duration of the CSP in general neurological and psychiatric disorders, we proceeded as follows:

- First, we used fully synthetic data that simulated simple MTS in order to gauge the adequacy of the proposed $wpIV$ measure. For that, a number of increasing levels of noise was used to contaminate the original data. This way, we should be able to assess if the proposed model and the corresponding index are capable of discovering the underlying true data without being affected by the added noise.
- Second, we generated and analyzed synthetic data that simulate the real EMG that are the ultimate target of the current thesis. These data were used with a two-fold objective: validating the previous results with artificial MTS, and comparing the performance of our

method with that of the current standard in the field, DM, in the presence of increasing levels of uninformative noise.

- Third: having validated the proposed $wpIV$ measure with artificial data, we proceeded to test it with real EMG data of control healthy patients (multiple trials per subject), who were not undergoing any type of motor rehabilitation, again comparing it with DM.
- Fourth: The duration of the CSP is likely to vary in pathological subjects undergoing rehabilitation. Our last set of experiments was aimed to provide the first preliminary evidence that the proposed $wpIV$ is a suitably robust method for the estimation of the CSP in patients undergoing rehabilitation from stroke.

These results are expected to be the starting point for research involving other types of neurological disorders.

5.1.4 The EMG data

For the experiments concerning human controls, 14 voluntary healthy right-handed subjects (8 women and 6 men, mean age ($\pm$ standard deviation) of 24.9($\pm$2.5) years) gave their consent to the study. They were informed about the experimental procedure and remained naive about the aim of the study. They had not been taking any drugs or alcohol during the previous day and during the day of the experiment.

TMS of the right M1 was produced using a biphasic Magstim Rapid 2 (Magstim Co., Withland, Dyfed, UK), with a 8-shaped coil (external diameter of each loop $9cm$). The coil handle pointed backwards and 45 degrees away from the midline. A tightly elastic cap marked with 1 cm $\times$ 1 cm left-side grid relative to the vertex (Cz) was fitted to each subject.

EMG signals were acquired using surface electrodes in a belly-tendon montage from the First Dorsal Interoseus (FDI) muscle of the right hand. The signal was amplified digitalized at $2Khz$. For each subject, the optimal hot spot for the right FDI was defined as the point of the grid where a MEP of $50\mu V$ was elicited with a probability of 50 $\%$ using the lowest stimulator output; that output was defined as the resting motor threshold (RMT). Afterwards, the output stimulator was set as 150 $\%$ of the resting motor threshold and 15 pulses were performed on the spot location with a controlled voluntary contraction of the right FDI muscle. Such voluntary contraction was set as the 10 $\%$ of the maximum voluntary contraction of the muscle, measured with a proper pressure gauge pinched.

A total of 15 trials per subject were recorded, and intervals between two consecutive TMS pulses were of at least 7 seconds, in order to avoid possible slow repetitive effects (See an example of the data in Fig. 5.2). Each signal covered an interval from 100 ms pre-stimulus to 500 ms post-stimulus. Subjects were not able to visually monitor the EMG signal, in order to avoid feedback effects due to the appearance of the signal.

For the experiments concerning pathological subjects, 8 chronic stroke patients (3 women and 5 men, mean age ($\pm$ standard deviation) of 64.9($\pm$10.5) years old) were identified and recruited by examining case-records from Bellvitge Hospital at Hospitalet de Llobregat, Barcelona, Spain. All patients had suffered from first-ever ischemic stroke causing upper limb weakness (3.5-4.5 on the Medical Research Council -MRC- scale). They were able to move the affected arm and the index finger without help from the healthy side.

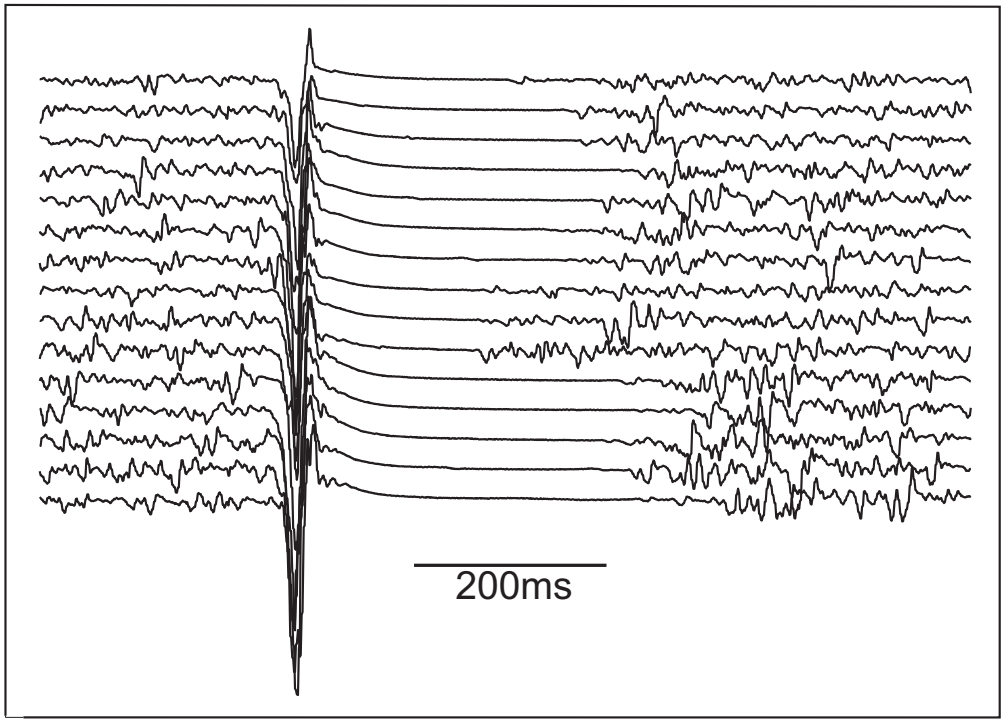

FIGURE 5.2: 15 trials of EMG signal acquired with surface electrodes in a belly-tendon montage from the FDI muscle of the right hand.

Exclusion criteria consisted of: 1) history of major psychiatric or previous neurological disease, including seizures and previous stroke; 2) cognitive impairment or history of dementia and 3) major comorbidity. All patients received multi-disciplinary post-stroke care appropriate to their clinical needs. The participants were all right handed. All the participants read and sign an informed consent prior to the participation on the protocol.

Over the course of four consecutive weeks, the patients received 20 individual Musical Supported Therapy (MST) sessions of 30 minutes each. Two different input devices were used to improve motor movements [63]: a MIDI-piano for fine motor movements and an electronic drum set comprising eight pads for gross motor movements. Drum pads (numbered from 1 to 8) were used to produce piano musical notes (C,D,E,F,G,A',B',C') rather than drum sounds. In a similar vein, The MIDI-piano was arranged in such a way that only 8 white keys (C,D,E,F,G,A',B',C') could be played by the patient. Each exercise was demonstrated by the therapist first and then repeated by the patient.

The patient moved through the therapy onto the next level of difficulty once he/she was able to complete the current level without errors. Thus, over the course of the therapy, patients proceeded from playing single notes to playing sequences of notes and beginnings of children's songs. Prior to, and at the end of the therapy course, patients were comprehensively evaluated by using TMS.

This TMS was performed using a focal air-cooled figure-8 coil (9 cm diameter each wing) attached to a biphasic Magstim Rapid 2 Stimulator. The choice of a biphasic rather than monophasic stimulator was based on availability of the stimulator and coils. The point of intersection of the figure-of-eight coil was placed against the skull and the coil was held at a 45 degrees angle to the sagittal with the handle oriented posteriolaterally [15].

A tightly elastic cap marked with 1 cm $\times$ 1 grid relative to the vertex (Cz) was fitted for each subject at baseline (first session) and reapplied consistently for second session to facilitate systematic sampling of scalp locations. Motor-evoked potentials (MEPs) were obtained from the First Dorsal Interosseus (FID) muscle of the hand contralateral to the stimulated hemisphere. Both affected and unaffected hemispheres were tested in each session.

At the begining of each session, a relatively high intensity of stimulation (typically 80% of maximum stimulator output) was used to locate the scalp position from which the MEPs of highest amplitude could be obtained and the location was marked on the cap to ensure consistent targeting of the *hot spot* location. The stimulator intensity was then dropped to 40% maximum stimulator output (MSP) and 10 pulses were delivered over the hotspot at a frequency of one every $5-10\ s$. If no MEP were observed in those 10 trials, the stimulator intensity was increased a 10% and 10 more trials were sampled. This procedure was repeated until the first evidence of a MEP appeared in the EMG trace. At this point, the stimulation threshold was determined by varying the stimulus intensity in 2% increments or decrements until the stimulus intensity was able to evoke 5 of 10 MEPs with an amplitude of at least $50\mu V$. This stimulus intensity is called the *resting motor threshold*. The active motor threshold (AMT) of the hot spot is defined as the stimulus intensity able to evoke 5 to 10 MEPs with an amplitude of at least $200\mu V$ at hot spot with sub-maximal shortening contraction.

This contraction was monitored visually for the subject using a pressure gauge pinched between the index finger and the thumb. The patient was instructed to maintain a 10% of his maximal voluntary contraction during each stimulation.

For the CSP assessment, 15 trials were recorded for each subject, and intervals between two consecutive TMS pulses were again of a duration of at least 7 seconds, to avoid possible slow repetitive effects. The trial time window was pulse-locked and the length of the EMG window was adapted to the patients, being the same for all of them.

5.1.5 Experiments and results

According to the experimental settings outlined above, the current section is structured in four parts. The first one reports the results corresponding to the experiments with synthetic general MTS. The second section reports the comparative results with synthetic data sets that emulate the typical EMG signal for the problem at hand. The third one reports results with real data of human controls. Finally, the last subsection offers preliminary results of experiments with real pathological subjects under rehabilitation from stroke.

5.1.5.1 Validation of the IV technique: experiments with synthetic data

The first set of experiments is meant to show the adequacy and usefulness of the $wpIV$ defined above. For that, we model a simple artificial set of MTS using both the standard GTM-TT and the VB-GTM-TT defined in the machine learning background chapter.

This basic data set consists of 3 time series built as a piecewise combination of step-like functions concatenating four periods of constant signal through three sudden transition change-points. The signal is contaminated with increasing levels of uninformative Gaussian noise (with standard deviations of, in turn: 0.01, 0.05 and 0.1 (see Fig.5.3, left hand-side column).

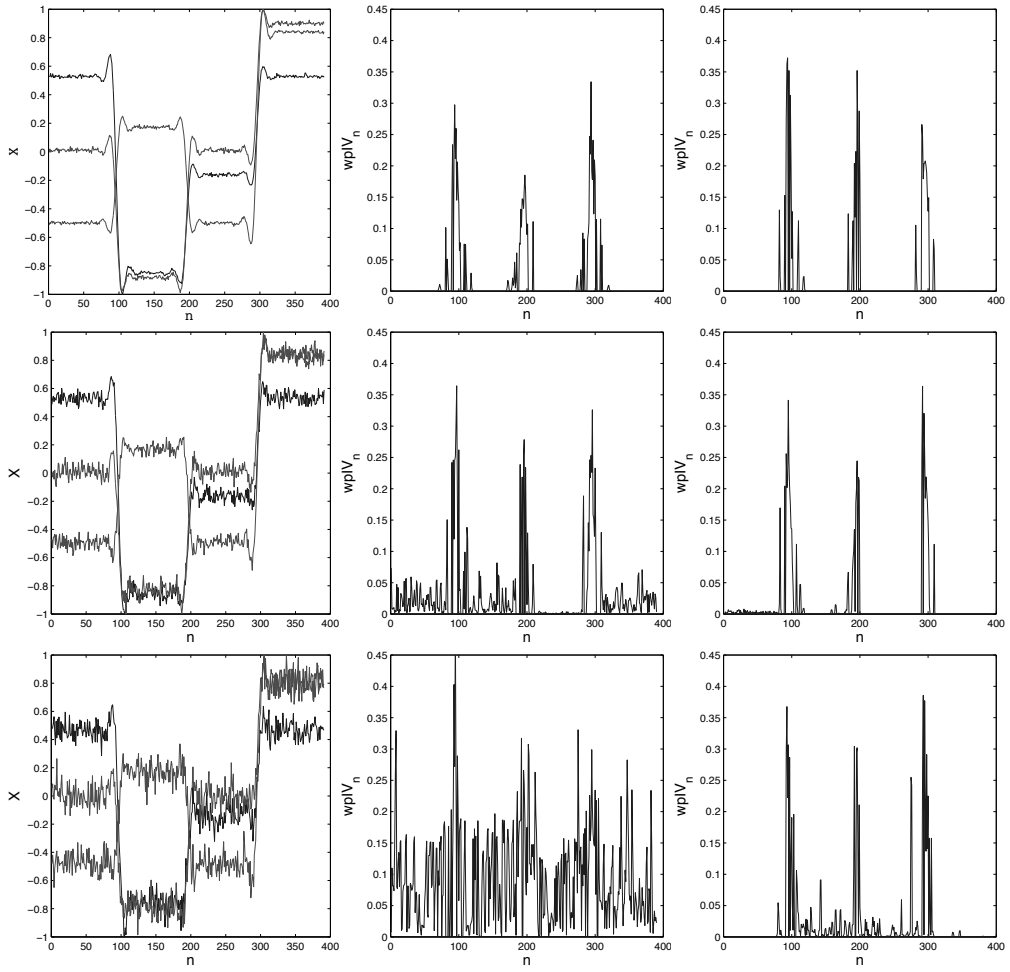

FIGURE 5.3: $wpIV$ for the artificial data at three noise levels of standard deviations: 0.01 (top row), 0.05 (middle) and 0.1 (bottom). The data are represented on the left column; the center column shows $wpIV$ results for GTM-TT; and the rightmost one, for VB-GTM-TT.

The $wpIV$ for both models and for the three noise levels is depicted in Fig.5.3 (center and right columns). At the lowest noise level (top row), the $wpIV$ corresponding to both models captures both the transitions and the periods of noise-related variability pretty well. At higher levels of noise, though, only the VB-GTM-TT (rightmost column) is able to keep faithfully modeling both of them. The $wpIV$ for the standard GTM-TT (center column), instead, clearly reveals that the model is overfitting the data, rendering the index useless for MTS segmentation through change-point detection.

This provides the first evidence of the robustness of the VB-GTM-TT model in the presence of noise. The standard GTM-TT overfits de MTS, while its alternative formulation within a Variational Bayesian framework is capable of retrieving the original underlying model with accuracy. This robustness should make the estimation of the $wpIV$ far more faithful than that rendered by any unregularized technique.

5.1.5.2 Experiments with EMG-like synthetic data

We now step forward in terms of experimental complexity, getting closer to the real EMG data that are the goal of the thesis. In this experiment, we compare the performances of DM and

VB-GTM-TT using artificial MEP-like data with added noise. We do so in order to gauge the strength of both methods when faced with noisy data.

The DM was implemented in Matlab (Mathworks) to identify the onset and the end of the CSP using mathematical criteria. The code involved a high-pass filtering process, as well as squaring and threshold detection. All trials were high-pass filtered (suppressing waveform activity below 50 Hz) to remove movement artifacts from the recordings.

The underlying artificial signal was piecewise-defined. Uninformative white noise was then generated and added to the CSP interval. The noise was used in three settings:

1. Varying the number of trials corrupted by noise, i.e., the ratio of noisy MTS in the dataset. Increasing levels of noise, with standard deviations of, in turn: 0.25, 0.5 and 1, were used.
2. Modifying the amplitude of the noise generated at the CSP, as represented by the standard deviation of the noise, with increasing values of 0.2, 0.4 and 0.6 (See Fig.5.4 for illustration).
3. Varying the number of time points corrupted by noise at the CSP, i.e., the ratio of noisy time points in the CSP, with values about 0.4, 0.7, 0.9 (See Fig.5.5).

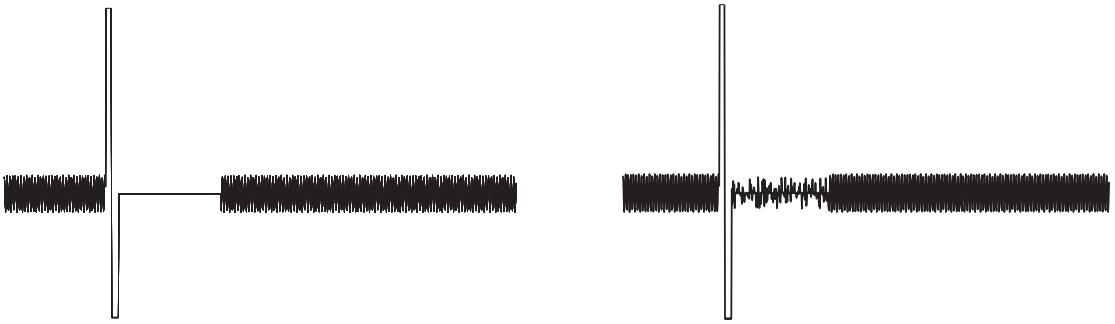

FIGURE 5.4: (Left) Artificial data created to simulate a typical CSP. (Right) The same artificial data with white noise added to the CSP interval.

Before and after the simulated CSP, the signal is defined as a high-frequency sinusoidal signal that emulates the voluntary muscle contraction that has to be performed by the subject.

A total of 27 noisy data sets were generated (Fig.5.5), taking into acoount the three possible settings described above. For each data set, 15 trials were generated and analyzed by DM and the VB-GTM-TT $wpIV$. In all data sets, the offset of silent period was homogeneously established at the point 426, in order to facilitate the analysis of accuracy measurement of both methods.

A full report of the results can be found in Table 5.1. It reveals that VB-GTM-TT was able to measure correctly the offset of the CSP in all contaminated data sets. However, DM was only able to measure correctly the offset when the noise amplitude was low. An illustration of this can be found in Fig.5.6. These are a very interesting results in two ways: First of all, VB-GTM-TT seems to be robust in the presence of noise, and, second, the weakness of DM seems to be more related to the amplitude of noise located on the silent period than to either the number of trials corrupted by noise, or to the ratio of corrupted points in each trial. In any case, since the VB-GTM-TT also correctly estimates the offset of the CSP for the noise-free data set, these results provide, overall, further evidence that the $wpIV$ calculated through VB-GTM-TT bears the potential of becoming a reliable tool for automating the determination of the CSP duration, specially in EMG recordings with high levels of noise.

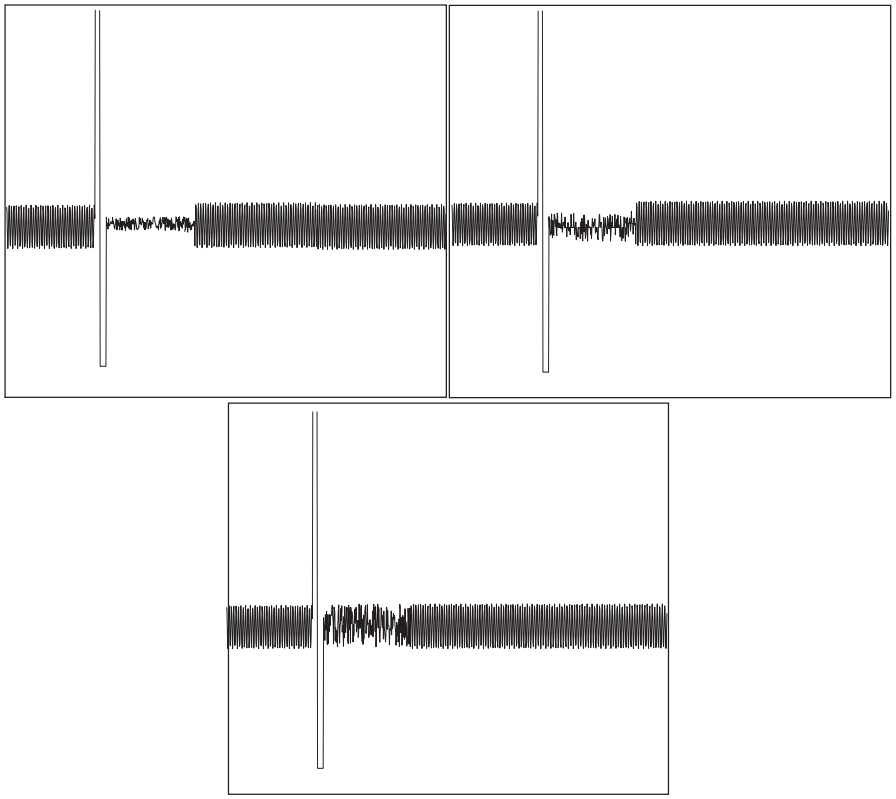

FIGURE 5.5: Three examples of synthetic data individual trials that simulate the real EMG, in the presence of increasing levels of uninformative noise. This time, in each one, the ratio of noisy points in the silent period is high (0.9), while the noise amplitude varies: 0.2 (top left), 0.4 (top right) and 0.6 (bottom).

5.1.5.3 Experiments with control subjects

We now progress to experimentation with real data acquired from human subjects. One of the main characteristics of this set of experiments is that we do not know any longer the true duration of the CSP. This means that there is no longer a true benchmark against which to compare the performance of the analyzed techniques.

The subjects involved in these experiments were voluntary controls, unaffected by the studied pathology and, therefore, not undergoing any therapy. These data were acquired with experimental purposes unrelated to any neurological injury.

Our hypothesis is that, since voluntary control data sets should in general not be contaminated by measurement noise, the estimations of the CSP duration yielded by DM and the VB-GTM-TT-based $wpIV$ will be quite similar, given that the accuracy of DM in these type of data has been tested in many experiments.

All trials for each subject were analyzed to determine the offset and the duration of the CSP, again using DM (with the previously described settings) and the VBGTM-TT-based $wpIV$. The squaring of the trace was employed to magnify and rectify the EMG activity as well as further isolate the MEP and the EMG flat period. The CSP onset was determined by the stimulus onset. The CSP offset was determined from the processed waveform and was the first point to exceed 25% of the mean pre-stimulus EMG amplitude within a generous window (between 100 and 500 ms after TMS stimulus) that consistently enclosed flat period after the end of the MEP.

TABLE 5.1: Mean CSP offset estimation for both models in all experimental settings described in the text.

Trials Corrupted	Points Corrupted	Noise Amplitude	DM	VBGTM-TT
25%	40%	0.2	426	426
25%	40%	0.4	394	426
25%	40%	0.6	401	426
25%	70%	0.2	426	426
25%	70%	0.4	393	426
25%	70%	0.6	417	426
25%	90%	0.2	426	426
25%	90%	0.4	401	426
25%	90%	0.6	417	426
50%	40%	0.2	426	426
50%	40%	0.4	378	426
50%	40%	0.6	345	426
50%	70%	0.2	426	426
50%	70%	0.4	360	426
50%	70%	0.6	392	426
50%	90%	0.2	426	426
50%	90%	0.4	360	426
50%	90%	0.6	351	426
100%	40%	0.2	426	426
100%	40%	0.4	306	426
100%	40%	0.6	303	426
100%	70%	0.2	426	426
100%	70%	0.4	302	426
100%	70%	0.6	301	426
100%	90%	0.2	426	426
100%	90%	0.4	303	426
100%	90%	0.6	301	426

All subjects completed the TMS protocol without difficulty. In almost all subjects, all trials passed a basic quality control. In 3 subjects, less than 3 trials were discarded from the analysis because of a contamination by movement artifact.

For illustration, we report next the results for two control subjects. Fig. 5.7 (left) shows their complete EMG. The corresponding estimation of the $wpIV$ is superimposed to Fig. 5.7 (center). The $wpIV$ provides a completely clean-cut delimitation of the CSP that allows the unambiguous estimation of its duration.

As explained in the introductory chapters, one of the unique advantages of a manifold learning method such as VB-GTM-TT is that it can provide a low-dimensional intuitive representation of high-dimensional MTS. Such visualizations in 2-D for the two controls used for illustration of the technique can be found in the right-hand side plot of Fig. 5.7. They show that the CSP is neatly represented by this model using separate states. This is a clean-cut indication that the change-points defining the onset and offset of the CSP are unambiguously detected and modeled by VB-GTM-TT.

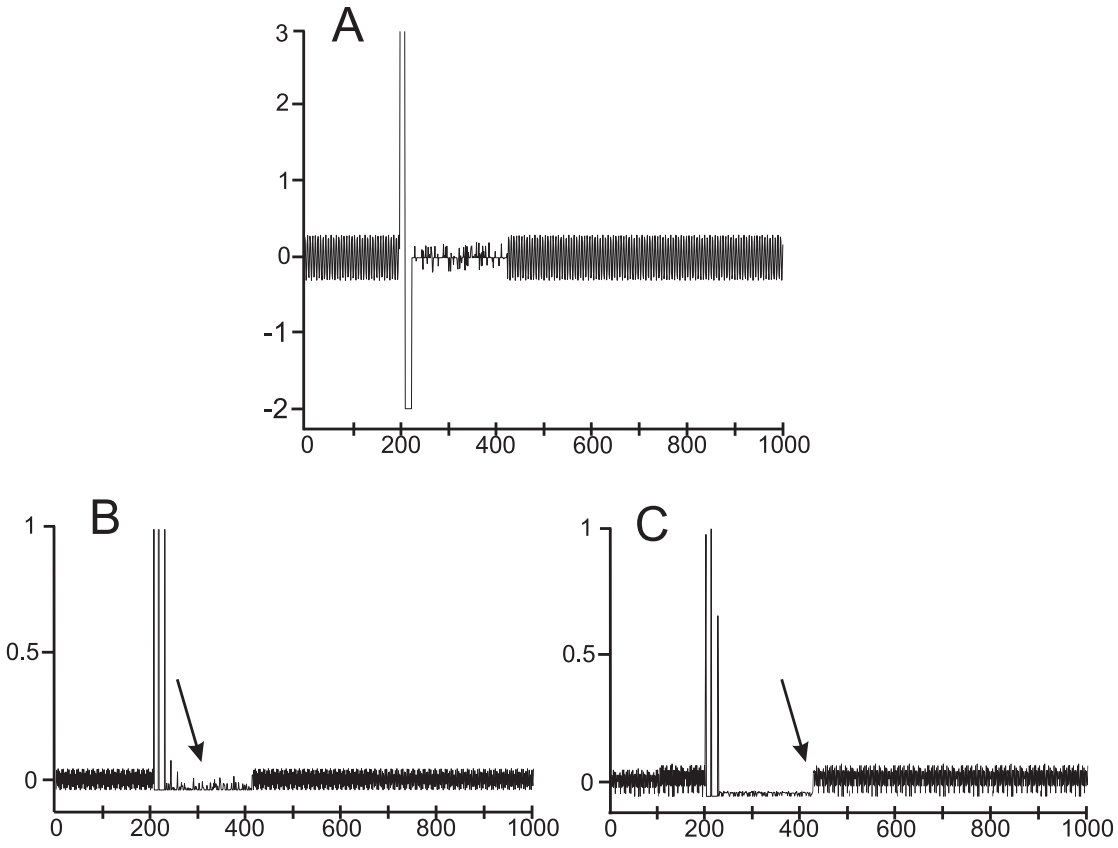

FIGURE 5.6: (A) Artificial data created to simulate a typical CSP pattern with Gaussian noise added to the CSP interval. (B) Normalized DM output after the squaring and filtering of the data represented in A. (C) Normalized output of the $wpIV$ obtained from the data represented in A. The arrows mark the estimations of the CSP offset in each method. DM clearly selects a wrong offset location.

A full report of the mean CSP durations of each subject obtained by both methods is provided in Table 5.2. It reveals that both methods estimate very similar durations of the CSP for all subjects but one, namely S14. This is a pretty result indeed for the following reasons: First of all, results overall corroborate our initial hypothesis: that for data (cases) likely not to be contaminated by measurement noise, such as the controls under study, there is not much differential advantage in using the VB-GTM-TT-based $wpIV$. In other words, these cases are simple enough for the CSP duration to be easily estimated by any reliable method. In any case, the similarity of the estimations indicate that the manifold learning method is yielding accurate estimations. Secondly, and taking our attention back to control S14, the data corresponding to this subject (see them represented in Fig.5.8), include many trials where the information about muscle contraction after the CSP was not properly acquired. As a result, the VB-GTM-TT-based method is not able to establish a coherent offset of the CSP.

5.1.5.4 Experiments with pathological subjects

The final stage of our experiments concerns the analysis of data corresponding to pathological patients who have undergone rehabilitation therapy. All subjects under study completed the TMS protocol without difficulty in two sessions (before and after undergoing musical therapy).

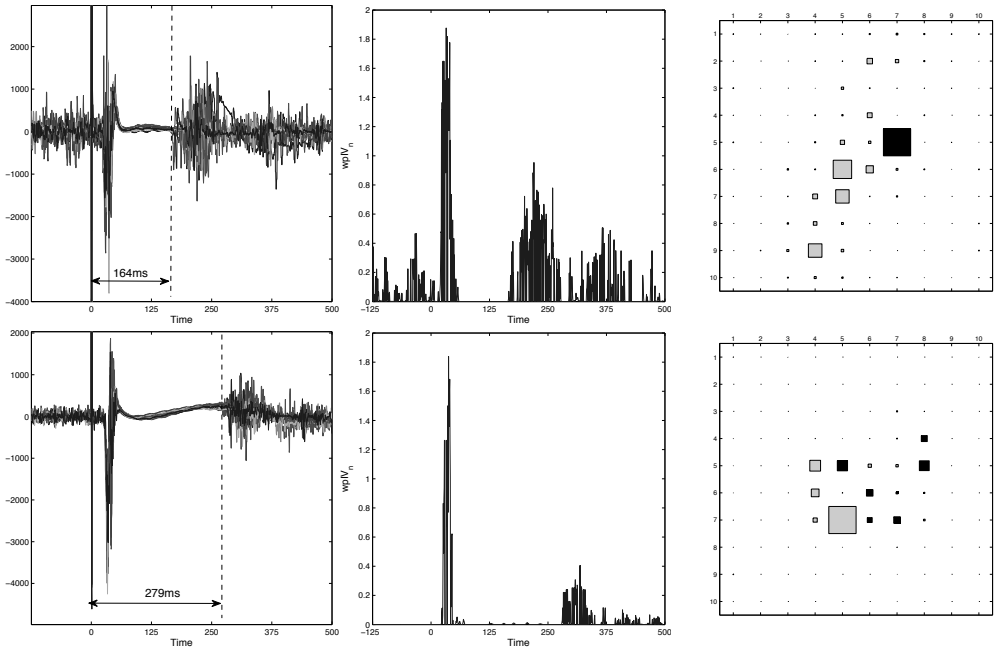

FIGURE 5.7: (Left) Visualization of the 15 EMG MTS corresponding to the complete sets of trials for two control subjects. Dashed lines delimit the CSP durations that were estimated using the $wpIV$. Middle) $wpIV$ for these subjects. Right) Visualization of the MTS in the VB-GTM-TT 2-D representation map. Squares represent model states and their size is an indication of the number of time points (as a ratio) assigned to each state. States filled in black correspond to the CSP.

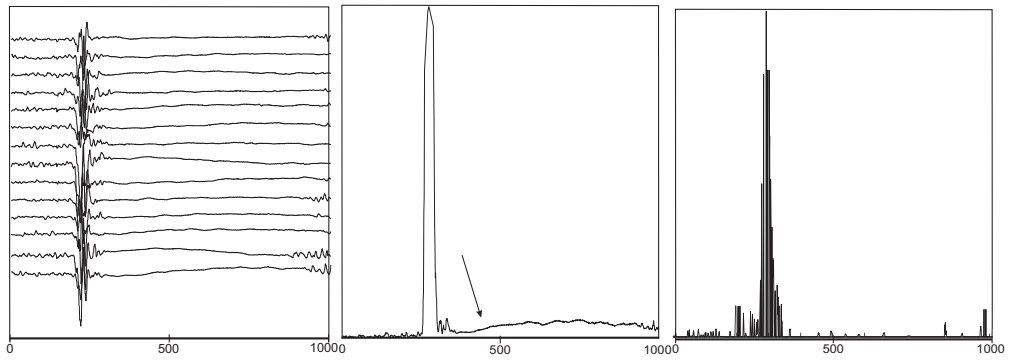

FIGURE 5.8: Left) The EMG time series for control subject S14. As can be observed, the return of voluntary contraction of the muscle was not acquired. Centre) Output of signal squaring by DM. The arrow marks the wrong offset of CSP estimated by DM. Right) $wpIV$ of the data; the method is not able to find any offset of the CSP.

However, although the pattern of EMG data acquired during voluntary muscle contraction between this population of subjects and the control subjects is quite similar, the acquisition conditions are very different. The patients in general present many difficulties to maintain a stable contraction of the muscle on the affected hand, and EMG spindles are registered during the CSP because of slow disruptions of inhibition processes. Also, the time necessary for the montage of the TMS and for data acquisition protocol is relatively high, causing in many cases discomfort and tiredness in the patients, making the task difficult. Given the circumstances, it seems probable that raw data will be contaminated by different kind of artifacts, making the the estimation of the CSP duration far more difficult than with control subjects.

TABLE 5.2: Table containing the mean of the estimated duration of the CSP for each control subject, using DM (central column) and VB-GTM-TT-based $wpIV$ (right hand-side column).

Subject	CSP DM (ms)	CSP GMT (ms)
S01	191	187
S02	283	284
S03	180	171
S04	259	263
S05	283	293
S06	260	261
S07	236	235
S08	236	239
S09	199	184
S10	288	277
S11	189	200
S12	258	245
S13	166	165
S14	157	203

TABLE 5.3: Table containing the mean of the CSP of each participant using DM (central column) and the VB-GTM-TT-based $wpIV$ (right hand-side column).

Subject	Session	DM (offset point)	VGTM-TT
P01	Pre	403	531
P01	Post	520	608
P02	Pre	622	668
P02	Post	652	515
P04	Pre	431	456
P04	Post	484	695
P05	Pre	666	615
P05	Post	443	519
P07	Pre	490	–
P07	Post	549	615
P09	Pre	658	704
P09	Post	435	520
P10	Pre	594	746
P10	Post	501	792
P11	Pre	301	–
P11	Post	433	653

The patients were evaluated twice. In a first session, before the starting of the neurorehabilitation sessions; and in a second session that was performed the day after the end of the rehabilitation procedures, one month after the first session. The original objective of this setting was to investigate the possible correlation between the shortening of the silent period after rehabilitation and the improvement in motor ability resulting from therapy.

Our hypothesis now (once both DM and VB-GTM-TT-based $wpIV$ have been evaluated with artificial data and real EMG from control subjects) is that VB-GTM-TT based $wpIV$ method will have a differential advantage in the measurement of the silent period in this data, because of the noisy nature of the data.

The mean CSP durations for each subject, obtained both by DM and the VB-GTM-TT-based $wpIV$ are listed in full in Table 5.3. In contrast to the results obtained with control subjects, the results obtained by both methods with stroke rehabilitation patients are very different, as we expected. It must be noted, though, that in two cases, the VB-GTM-TT-based $wpIV$ was not able to establish the offset of the SP (Fig.5.9).

Results in Table 5.3 do suggest that there is no obvious correlation between the shortening of the silent period after rehabilitation and the improvement in motor ability resulting from therapy. There are in fact more subjects that show an increase in the duration of the CSP after therapy.

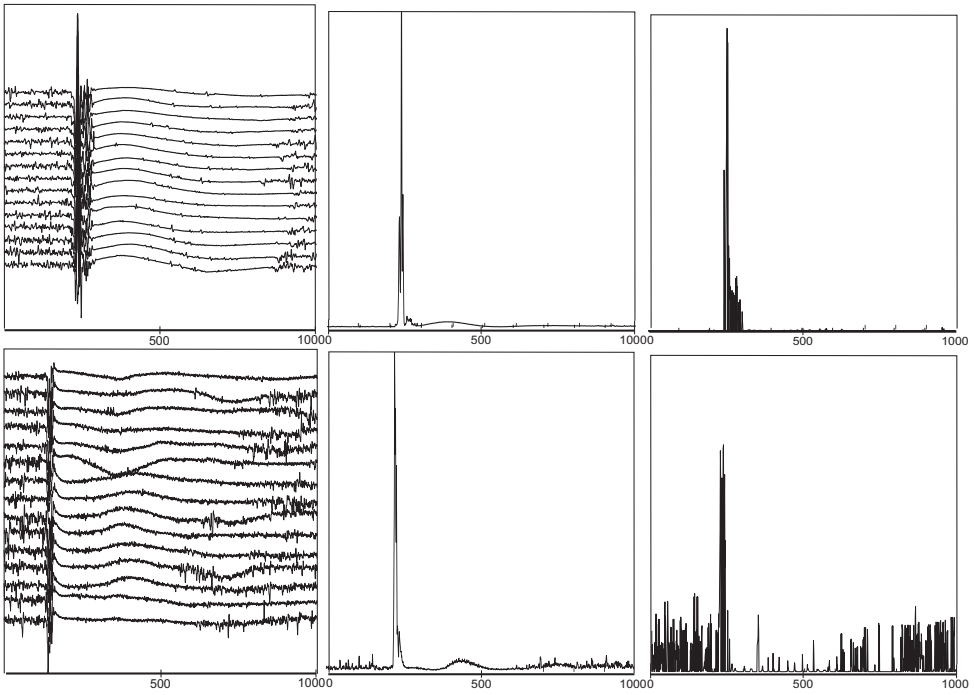

FIGURE 5.9: Left) The EMG for subjects P011 and P07 (pre-therapy). It is clear that for the P07 patient (top), the return of voluntary contraction of the muscle was not properly acquired in several trials. As a result, the VB-GTM-TT-based method was not able to measure the offset of the CSP.

In the same way we did it for control subjects, illustrative results are presented in Fig.5.10 for one stroke patient. All data trials and the estimation of the CSP duration is shown in the left hand-side, the calculated $wpIV$ over time is shown in the centre, while the low-dimensional VB-GTM-TT representation of the data is shown on the right hand-side.

Results for measurements before rehabilitation are presented in the top row, whereas results after rehabilitation are presented in the bottom row. The $wpIV$ provides a clean-cut delimitation of the CSP that allows the unambiguous estimation of its duration. The 2-D visualization shows that the CSP is accordingly described almost in full by separate model states. In this patient at least, the CSP duration is shown to have considerably shortened due to rehabilitation, as we originally expected.

Let us also illustrate the difference of applying either the proposed VB-GTM-TT-based method or DM to two stroke patients. Results are shown in Fig.5.11: raw data for all trials (left); results of the application of DM (center) and of VB-GTM-TT (right). In both cases, the DM estimation of the duration of the CSP places the offset time of the silent period in a noise-affected location

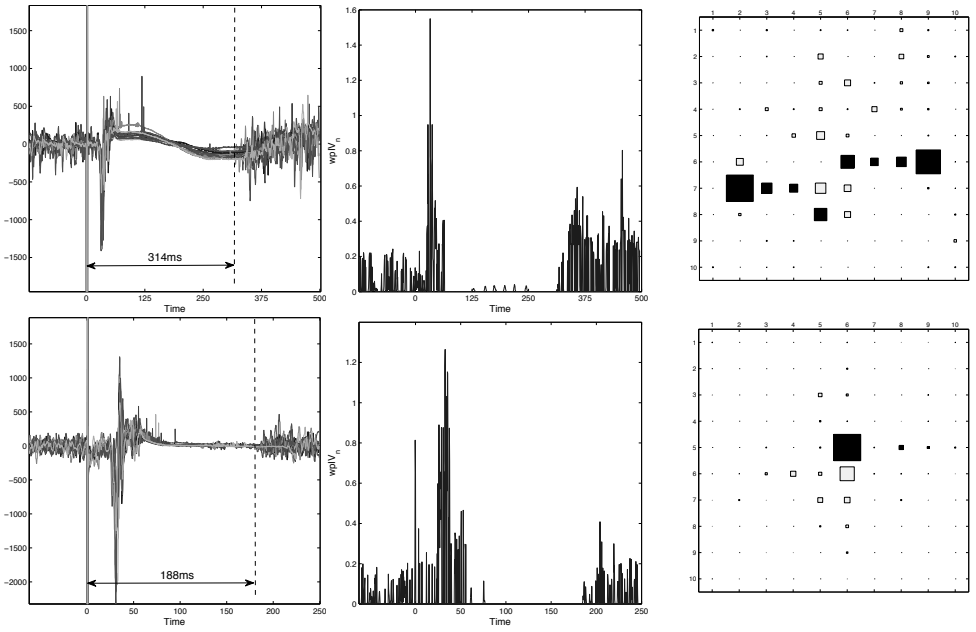

FIGURE 5.10: Stroke rehabilitation patient. Top) Before rehabilitation; bottom) after rehabilitation. Representation as in the previous figure.

of the refractory period. However, also in both cases our proposed method marks the end of the silent period in a correct way, which is coherent with the visual output of the method.

These results confirm that Daskalakis method does not seem to be a reliable tool to measure the SP in EMG acquisition from patients, probably because of the huge quantity of noise present on the data, and also suggest that VB-GTM-TT-based $wpIV$ could be a strong-to-noise method able to measure SP in data acquired from patients.

5.1.6 Conclusion

The CSP is known to be a key feature of EMG recordings from subjects affected by many neurological and psychiatric disorders. The CSP in such patients is usually abnormal, reflecting altered cortical inhibition. For this reason, the accurate determination of the duration of the CSP is a worthy research target from a clinical point of view.

Traditionally, the CSP has been calculated using *ad hoc* techniques prone to subjective variability. Only recently, computer-based automated procedures for the estimation of its duration have become mainstream. The current standard procedures, though, are limited mainly for their lack of robustness in the presence of noise. In this chapter, we have put forward a novel method for the analysis of the CSP. It is based on the robust detection of change points in the EMG MTS using a Statistical Machine Learning method of the manifold learning family. This model is meant to reveal the underlying structure of the MTS signal even if contaminated by uninformative noise.

To test this proposed new technique, we have gone in this chapter through a battery of experiments. Let us recall them in brief:

- First, we used fully synthetic data that simulated simple MTS in order to gauge the adequacy of the proposed index of variability based on VB-GTM-TT. A number of increasing

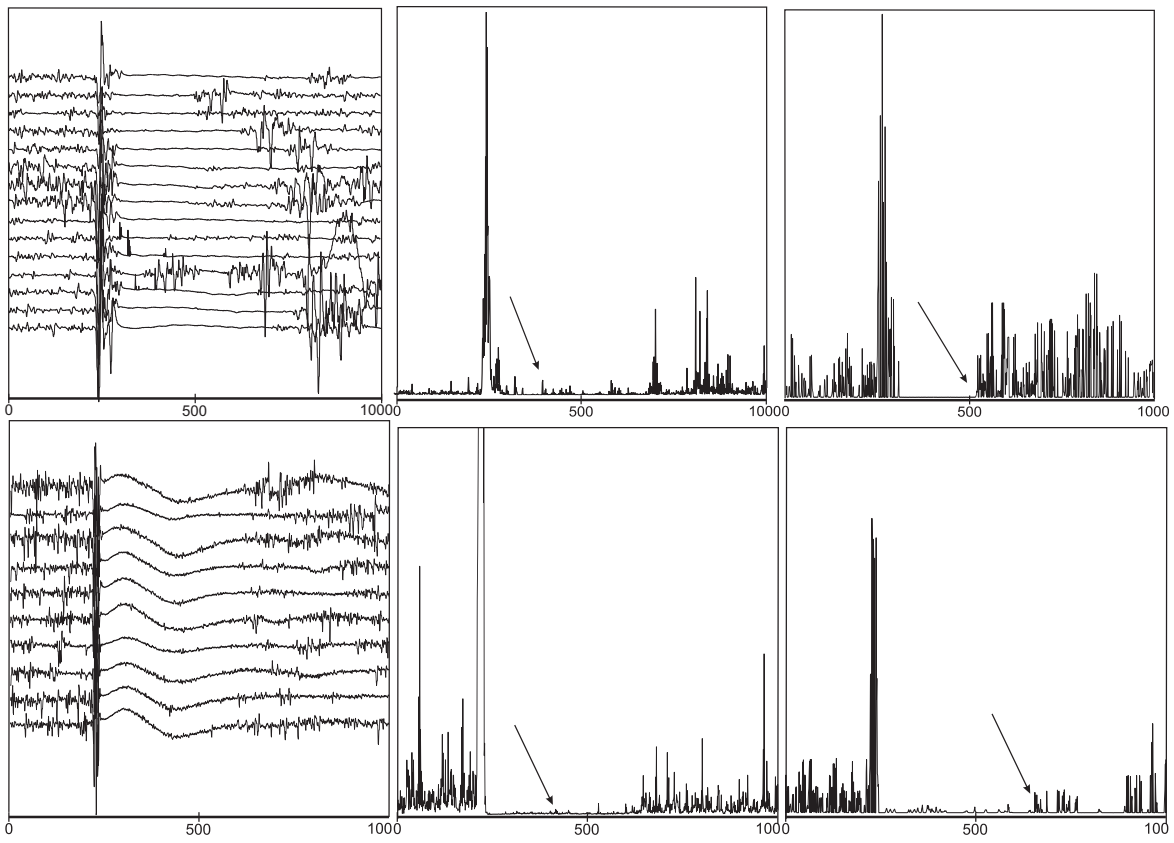

FIGURE 5.11: Comparison between DM and $wpIV$ in two stroke patients. Left) 15 EMG trials acquired maintaining voluntary contraction of the muscle. Center) Signal after application of DM. Right) $wpIV$ results. The arrow marks the estimation of the CSP offset in each method.

levels of noise were used to contaminate the MTS. The model showed that it was capable of discovering the underlying clean data without being too badly affected by the noise. These findings provided the first evidence of the adequacy and usefulness of the technique.

- We then tried to replicate these results using synthetic data that were closer to the real EMG that is the object of the thesis. Importantly, we also compared the performance of the VB-GTM-TT-based method with the DM, which is the current standard in the field. Increasing levels of uninformative noise were added in several settings. The results overall showed that VB-GTM-TT was consistently more robust and accurate than CSP in the estimation of the CSP. Also, importantly, the results suggest that the weakness of the standard DM is more related to the amplitude of the noise located on the silent period than to either the number of trials affected by noise, or the number of corrupted points in each trial. All this provided further evidence that the $wpIV$ may become a reliable tool for automating the determination of the CSP duration, specially in EMG recordings with high levels of noise.

- Real EMG data of control healthy subjects were subsequently used to test the new proposed method. This data is supposed to be almost noise- and acquisition artifacts- free. For this reason, we hypothesized that both compared methods should perform similarly. The results obtained were very similar for most subjects, corroborating the hypothesis. The only subject for which some data trials were not properly acquired was the only exception.

- We were finally in the position to confidently experiment with real EMG data acquired from stroke patients. It has been suggested that the duration of the CSP should naturally vary in pathological subjects undergoing rehabilitation, indicating the progrees due to therapies aplied to the patients. Both methods were applied to measure CSP duration on data recorded both before and after the rehabilitation process. The data were expected to be contaminated by noise and artifacts because of the acquisition conditions. Results provided some important evidence: Firstly, that the measuremets using both methods were again very different from each other, as it could have been expected. Since the results obtained by DM are not coherent with what the data suggest through direct visualization, these experiments support our hypothesis that the $wpIV$ measure behaves robustly in the presence of noise, while DM fails to provide an appropriate estimation. Secondly, and contradicting our hypothesis, there was preliminary evidence of no clear correlation between the dynamics of silent periods and the performance of this therapy.

Chapter 6

Summary and Conclusions

6.1 Summary and Conclusions

Over the last decade, biomedicine has become a data-intensive field of research in which new data acquisition techniques appear at a staggering pace. The increasing reliance on microarrays data in genomics, and on protein chips and tissue arrays in proteonomics, add to the wealth of information about active metabolic pathways that is available also from many other non-invasive measurement techniques. The challenge of managing the complexity of biomedical data invites us to go one step further than traditional statistics and resort to advanced techniques that are capable to take the best out of both machine learning and statistics. We are to find such techniques in the area of statistical machine learning.

Computational neuroscience has reached, in parallel to biomedicine, a stage of quick development driven by the new data generating possibilities of new noninvasive measurement techniques. In this development, it has gathered strength from fields as diverse as systems biology, genetics, bioinformatics, signal processing, artificial intelligence, and psychology.

Brain processes can be studied from physiological data. This thesis has been concerned with the analysis of rehabilitation of patients who suffered a brain stroke. This neuroscience problem is in the frontier with clinical medicine. The motor rehabilitation of patients affected by this pathology can be explored through the EMG signal elicited in controlled motor tasks. EMG data usually have a strong component of noise resulting from measurement conditions that must be processed using robust procedures in order to extract any kind of medically usable knowledge. This is the case for the purpose of both medical research and medical practice.

The thesis has explored in some detail one of the key parameters of this type of EMG signal: the CSP, elicited by TMS. The existing CSP measuring methods are yet imprecise and known to yield a significant error due to their sensitivity to noise. Yet, this parameter is known to be very useful for the diagnosis and prognosis of motor progress in patients that have suffered brain stroke. The correct estimation of the CSP duration is difficult because of the presence of uninformative noise in the measured signal.

We have shown, through a battery of experiments, the usefulness of a manifold-constrained Hidden Markov Model formulated within a variational Bayesian framework in dealing appropriately with the problem of CSP duration estimation. The proposed model is imbued with regularization properties that minimize the negative effect of the presence of noise in the analyzed EMG

data.We hypothesized that this method would be a reliable tool to calculate the offset of the refractory period even in the presence of relatively large amounts of noise. The experiments have provided evidence to support such hipothesis. They have also provided evidence of the superiority of the proposed method with respect to the existing standard.

The contributions of this master thesis are briefly summarized as follows:

We have developed a technique based on a Bayesian Variational variation of the GTM-TT model that detects sudden transitions and change points on EMG data, with the aim to identifying the CSP as a time subsequence of little variability, even with presence of noise. An Index of Variability (namely the weighted-prototype IV) has been described: it allows measuring the degree of variability of a subsequence. It naturally takes advantage of the data regularization properties of the underlying model and true change-points can be clearly detected even with presence of noise. Using a full experimental setting, that included artificial data sets with increasing levels of noise and real EMG data acquired from control subjects and patients, evidence has been gathered to support two important findings. Firstly, the wpIV has been shown to measure the CSP duration with better accuracy that the existing standard methods in use in the presence of noise, while matching the performance of the existing standard in noise-free data. Secondly, we have not found enough evidence to support the expected correlation between the shortening of the CSP and the performance of the rehabilitation therapy undergone by the pathological patients. Some studies have reported that a shortening of the CSP is the result of cortical reorganization on the affected hemisphere on stroke patients, and is related to motor improvement. Nevertheless, such studies used data in which different therapies, such that constraint-induced therapy [46], were used. Constraint-induced therapy is far more agressive on the patient than MST (which was the one applied to the patients analyzed in this study), in the sense that it involves more direct stimulation of cortical circuitries through rehabilitation, which means that the effects of the induced plasticity are likely to be more immediate. Further tests with a larger number of patients would be required to provide us with a clearer picture of the possible correlation between the shortening of the CSP and the performance of the MST.

6.2 Future directions

6.2.1 Clinical studies

In this thesis, the proposed method has been applied to data acquired from stroke patients that have undergone a rehabilitation process, but many other clinical fields are currently interested in the study of inhibition phenomena. For example, in focal dystonia, a neurological condition affecting a muscle or group of muscles in a part of the body causing an undesirable muscular contraction or twisting, studies of inhibition are crucial to understand its dynamics and prognosis. However, there are many problems affecting the measurement of cortical silent periods [35], because the muscle control problems affect the EMG data acquired from these patients.

Also, inhibition is a parameter related with the diagnosis and prognosis of Parkinson and Huntington diseases[40], which are the cause of tremoring and loss of control of muscles difficulting the CSP measurement.

Future could be devoted to the application of our proposed model to other neurological conditions as those outlined.

6.2.2 Application to other EMG parameters

We have focused on the application of VB-GTM-TT to the measurement of CSP. For further validation of this model as a tool to analyze EMG data, we should test it in the calculation of other inhibitory and excitatory parameters. An interesting parameter where our method could be applied are MEP latencies, i.e., the time elapsed between the TMS pulse onset and the MEP. This could also be treated as a change-point detection problem, since the MEP entails a sudden change in the EMG signal. The behavior of these latencies is also related to many neurological disorders of the central neural system and peripheral nerves, and no automatic methods are currently used to measure it. Double-paired stimulation studies could also be enriched with the application of our model.These studies involve performing two consecutive pulses in different (or equal) magnetic stimulator output (percentage of stimulation), and measuring the effects of the first stimulation on the second one. Depending on how this two stimulations are performed, and also depending on how much time separates both stimulations (generally, less than 50ms), excitatory and inhibitory effects will be generated, reflected on the second MEP acquired using EMG. Our method could be used to estimate both the latency between stimulations and also the amplitude of both MEP activations; processes which are currently performed by direct visualization of the data.

6.2.3 Application on other Neuroimaging Techniques

Further application of our model could extend to other neuroimaging techniques, such as EEG or functional Magnetic Resonance Imaging (fMRI).

EEG is the recording of electrical activity long the scalp produced by the firing of neurons within the brain. This signal seems to be very sensible to ocular movements and muscular artifacts. Currently, there are many ocular artifact rejection algorithms in use, some of them based on Independent Component Analysis and Artificial Neural Networks paradigms. Nevertheless, this yet remains an open problem. Our method could be useful in the automatic detection of artifacts on EEG recordings, since we have provided evidence that it is a powerful method of change-point detection and resilient-to-noise visualization. Another application in the field of EEG research could be in the study of event-related potentials (ERP).

An ERP is any measured brain response that is directly the result of a thought or perception. The study of ERP involves the repetition of some cognitive task, such as movements in response to any perceptual stimuli. The ERP is the result of averaging the activity of the brain when the task is performed. Some of these studies require the performance and registration of many trials (more than 100 in some cases), in order to obtain a clear and usable ERP. This makes data acquisition difficult for some subjects, such as pathological patients. Our method could be used as a change-point analyzer in raw EEG data, reducing the time required to register the data. Our method could also be used in this field as a basis for the future development of an on-line Brain Computer Interface (BCI) based on hand movement detection mediated by EEG.Further application of our model could extend to other neuroimaging techniques, such as EEG or functional Magnetic Resonance Imaging (fMRI). EEG is the recording of electrical activity long the scalp produced by the firing of neurons within the brain. This signal seems to be very sensible to ocular movements and muscular artifacts. Currently, there are many ocular artifact rejection algorithms in use, some of them based on Independent Component Analysis and Artificial Neural Networks paradigms. Nevertheless, this yet remains an open problem. Our method could be useful in the automatic detection of artifacts on EEG recordings, since we have

provided evidence that it is a powerful method of change-point detection and resilient-to-noise visualization. Another application in the field of EEG research could be in the study of event-related potentials (ERP). An ERP is any measured brain response that is directly the result of a thought or perception. The study of ERP involves the repetition of some cognitive task, such as movements in response to any perceptual stimuli. The ERP is the result of averaging the activity of the brain when the task is performed. Some of these studies require the performance and registration of many trials (more than 100 in some cases), in order to obtain a clear and usable ERP. This makes data acquisition difficult for some subjects, such as pathological patients. Our method could be used as a change-point analyzer in raw EEG data, reducing the time required to register the data. Our method could also be used in this field as a basis for the future development of an on-line Brain Computer Interface (BCI) based on hand movement detection mediated by EEG.

fMRI measures the hemodynamic response (change in blood flow) related to neural activity in the brain or spinal cord of humans or other animals, through changes of a powerful magnetic field. It generates an image with a very precise maximum spatial resolution of 0.5mm of the activations of the whole brain elicited by a particular cognitive task. As opposed to EEG, this technique has a poor time resolution, because of the difficulty of measuring the hemodynamic signal. Our method could be tested in such data, in order to allow high-frequency image acquisitions and calculate more accurately the hemodynamic changes related to a cognitive task. Also, this signal is very sensitive to motion, and current algorithms of normalization of the data are affected by the presence of noise in the data. Our Index of Variability could be used also to correct the images contaminated by noise and allow a more accurate analysis of this signal. To this aim, some adaptations of our method would be required.

fMRI measures the hemodynamic response (change in blood flow) related to neural activity in the brain or spinal cord of humans or other animals, through changes of a powerful magnetic field. It generates an image with a very precise maximum spatial resolution of 0.5mm of the activations of the whole brain elicited by a particular cognitive task. As opposed to EEG, this technique has a poor time resolution, because of the difficulty of measuring the hemodynamic signal. Our method could be tested in such data, in order to allow high-frequency image acquisitions and calculate more accurately the hemodynamic changes related to a cognitive task. Also, this signal is very sensitive to motion, and current algorithms of normalization of the data are affected by the presence of noise in the data. Our Index of Variability could be used also to correct the images contaminated by noise and allow a more accurate analysis of this signal. To this aim, some adaptations of our method would be required.

6.3 Appendix 1

Two publications are the result of this thesis

- Olier, I., Amengual, J. and Vellido, A. Segmentation of EMG time series using a variational Bayesian approach for the robust estimation of cortical silent periods. In Procs. of the 18th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN), Bruges, 2010.
- Olier, I., Amengual, J. and Vellido, A. A variational Bayesian approach for the robust estimation of cortical silent periods from EMG time series of brain stroke patients. Invited for publication, *Neurocomputing*, 2011.

6.4 Appendix 2: List of acronyms

- CSP: Cortical Silent Period.
- EEG: Electroencephalography.
- EMG: Electromiography.
- ERP: Event Related Potential.
- fMRI: Functional Magnetic Resonance Imaging.
- GABA: Gamma-aminobutiric acid.
- GTM: Generative Topographic Maps.
- GTM-TT: Generative Topographic Maps-Through Time.
- IV: Index of Variability.
- MEG: Magnetoencephalography.
- MEP: Motor Evoked Potential.
- MST: Musical Supported Therapy.
- MTS: Multivariate Temporal Series.
- PET: Positron Emision Tomography.
- RTM: Resting Motor Threshold.
- TMS: Trasncranial Magnetic Stimulation.
- VB-GTM-TT: Variational Bayesian Generative Topographic Maps-Through TIme.
- $wpIV$: weighted-prototype Index of Variability.

Bibliography

[1] Adkins, D.: Motor training induces experience-specific patterns of plasticity across motor cortex and spinal cord. J Appl Physiol **99**, 1558–1586 (2006)

[2] Bangert, M., Altenmueller, E.: Mapping perception to action in piano practice: A longitudinal dc-eeg-study. BMC Neuroscience **4**, 26–36 (2003)

[3] Bangert, M., Peschel, T., Schlaug, G., Rotte, M., Drescher, D., Hinrichs, H.: Shared networks for audotpry and motor processing in professional pianists: Evindence from fmri conjunction. Neuroimage **30**, 917–926 (2006)

[4] Barker, A., Freeston, I., Jalinous, R.: Non-invasive stimulation of human motor cortex. Lancet **2**, 1106–1107 (1985)

[5] Bartholomew, D.J.: Latent Variable Models and Factor Analysis. Charles Griffin and Co. Ltd, London (1987)

[6] Bartholow, R.: Experimental investigations into the function of the human brain. American Journal of the Medicas Sciences **134**(1), 205–313 (1874)

[7] Baum, L., Egon, J.: An inequality with applications to statistical estimation for probabilistic functions for a Markov process and to a model for ecology. B. Am. Meteorol. Soc. **73**, 360–363 (1967)

[8] Beal, M.: Variational algorithms for approximate Bayesian inference. Ph.D. thesis, The Gatsby Computational Neuroscience Unit, Univ. College London (2003)

[9] Berardelli, A.: Transcranial magnetic stimulation in movement disorders. Clinical Neurophysiology **51**, 276–280 (1999)

[10] Bishop, C.M., Hinton, G., Strachan, I.: GTM Through Time. In: IEE Fifth Int. Conf. on Artif. Neural Net., Cambridge, U.K., pp. 111–116 (1997)

[11] Bishop, C.M., Svensén, M., Williams, C.K.I.: Developments of the Generative Topographic Mapping. Neurocomputing **21**(1–3), 203–224 (1998)

[12] Bishop, C.M., Svensén, M., Williams, C.K.I.: GTM: The Generative Topographic Mapping. Neural Comput. **10**(1), 215–234 (1998)

[13] Bishop, C.M., Tipping, M.E.: A hierarchical latent variable model for data visualization. IEEE Trans. Pattern Anal. Mach. Intell. **20**(3), 281–293 (1998)

[14] Bizzi, E.: New perspectives on spinal motor systems. Naute Re vNeurosc **1**, 101–108 (2000)

[15] Brasil-Neto, J., Pascual-Leone, A., Valls-Sole, J., Cohen, L., Hallet, M.: Focal trasncranial magnetic stimulation and response bias in a forced-choice task. J Neurol Neurosurg Psychiatry **55**(10), 964–966 (1992)

[16] Byrnes, M., Thickbroom, G., Phillips, B., Mastaglia, F.: Long-term changes in motor cortical organization after recovery from subcortical stroke. Brain Research **886**, 278–287 (2001)

[17] Caramia, M., Palmieri, M., Desiato, M.: Pharmacologic reversal of cortical hyperexcitability in patients with als. Neurology **54**, 58–64 (2000)

[18] Chatfield, C.: Time Series Forecasting. Chapman & Hall/CRC Press (2000)

[19] Chen, R., Lozano, A., Ashby, P.: Mechanism of the silent period following transcranial magnetic stimulation: evidence from epidural recordings. Exp Brain Research **128**, 539–542 (1999)

[20] Cicinelli, P., Traversa, R., Rossini, P.: Post-stroke reorganization of brain motor output to the hand: a 2-4 month follow-up with focal magnetic transcranial magnetic stimulation. Electroencephalogr Clin Neurophysiol **105**, 438–450 (1997)

[21] Classen, J., Schnitzler, A., Bonjofski., F.: The motor syndrome associated with exaggerated inhibition within the primary motor cortex of patients with hemiparetic. Brain **120**, 605–619 (1997)

[22] Cortes, C., Vapnik, V.: Support-vector networks. Machine Learning **20**, 273–297 (1995)

[23] Daskalakis, Z., Molnar, G., Christensen, B., Sailer, A., Fitzgerald, T., Chen, R.: An automated method to determine the transcranial magnetic stimulation-induced contralateral silent period. Clinical Neurophysiology **114**, 938–944 (2003)

[24] Dobkin, B.: Rehabilitation after stroke. New Engl J Med **352**, 1677–1684 (2005)

[25] Dobkin, B.: Training and exercise to drive poststroke recovery. Nature Rev **4**(2), 76–85 (2007)

[26] Ferrier, D.: Experimental researches in cerebral physiology and pathology. West Riding Lunatic Asylum Medical Reports **3**, 1–50 (1873)

[27] Foltys, H., Krings, T., Meister, I., Sparing, R., Topper, R.: Motor representation in patients rapidly recovering after stroke: a functional magnetic resonance imaging and transcranial magnetic stimulation study. Clin Neurophysiology **114**, 2404–2415 (2003)

[28] Fregni, F., Boggio, P., Mansur, C., Leone, A.P.: Transcranial direc current stimulation of the unaffected hemisphere in stroke patients. Neuroreport **16**, 1551–1555 (2005)

[29] Fritsch, G., Hitzig., E.: Uber die elektrische erregbarkeit dess frosshirns. translation by g. von bonin. The Cerebral Cortex **1**, 73–96 (1870)

[30] Garvey, M., Ziemann, U., Becker, D., Barker, C., Bartko, J.: New graphical method to measure silent periods evoked by transcranial magnetic stimulation. Clin neurophysisology **14**, 1451–1460 (2001)

[31] Girolami, M.: Latent variable models for the topographic organisation of discrete and strictly positive data. Neurocomp. **48**, 185–198 (2002)

[32] Gualtierotti, T., Paternson, A.S.: Electrical stimulation of the unexposed cerebral cortex. Journal of Physiology **125**, 278–291 (1954)

[33] Hallet, M.: Transcranial magnetic stimulation: A primer. Neuron **55**(1), 187–199 (2007)

[34] Hastie, T.: Principal curves and surfaces. Tech. rep., Department of Statistics, Stanford Univ. (1984)

[35] Huang, Y., Rothwell, J., Lu, C., Chen, R.: Restoration of motor inhibition through an abnormal premotor-motor connection in dystonia. Mov Disord **6**, 689–696 (2010)

[36] Jolliffe, I.T.: Principal Component Analysis. Springer-Verlag, New York (1986)

[37] Jones, T., Schallert, T.: Use-dependent growth of pyramidal neurons after neocortical damage. J Neurosci **14**, 2140–2152 (19944)

[38] Kabán, A., Girolami, M.: A dynamic probabilistic model to visualise topic evolution in text streams. J. Intell. Inf. Syst. **18**(2–3), 107–125 (2002)

[39] Kobayashi, M., Pascual-Leone, A.: Transcranial magnetic stimulation in neurology. Lance Neurotogy **2**, 145–56 (2003)

[40] Koch, G., Brusa, L., Stanzione, P.: Cerebellar magnetic stimulation decreases levodopa-induced dyskinesias in parkinson disease. Neurology **72**, 113–119 (2009)

[41] Kohonen, T.: Self-Organizing Maps (3rd ed). Springer-Verlag, Berlin (2001)

[42] Kramer, M.A.: Nonlinear Principal Components Analysis using Autoassociative Neural Networks. AIChe Journal **37**(2), 233–243 (1991)

[43] Kwakkel, G.: Effects of augmented exercise therapy time after stroke: a meta-analysis. Stroke **35**, 2529–2539 (2004)

[44] Lawley, D.N., Maxwell, A.E.: Factor Analysis as a Statistical Method. 2 edn. Butterworth & Co., London (1971)

[45] Liepert, J.: Motor cortex disinhibition of the unaffected hemisphere after acute stroke. Muscle Nerve (23), 501–519 (2001)

[46] Liepert, J.: Motor cortex excitability in stroke before and after constraint-induced movement therapy. Cogn Behav Neurol **19**(1), 41–47 (2006)

[47] Liepert, J., Haevernick, K., Weiller, C., Barzel, A.: The surround inhibition determines therapy-induced cortical reorganization. Neuroimage **32**(3), 1216–1220 (2006)

[48] Mardia, K., Kent, J., Bibby, M.: Home Health Handbook. Merck Press (2007)

[49] Matsuzaki, M.: Factors critical for the plasticity of dendritic spines and memory storage. Neurosci Res **57**, 1–9 (2007)

[50] Merton, P., Morton, H.: Stimulation of the cerebral cortex in the intact human subject. Nature **285**(287), 115–147 (1980)

[51] Meyer, M., Elmer, S., Bauman, S., Jaencke, L.: Short-term plasticity in the auditory system: Differential neural responses to perception and imagery of speech and music. Restoration, Neurology and Neurosciences **25**, 411–431 (2007)

[52] Olier, I.: Variational bayesian algorithms for the generative topographic mapping and its extensions. Ph.D. thesis, Universitat Politecnica de Catalunya (2008)

[53] Olier, I., Vellido, A.: Comparative assessment of the robustness of missing data imputation through Generative Topographic Mapping. In: Lect. Notes Comput. Sc., vol. 3512, pp. 787–794 (2005)

[54] Olier, I., Vellido, A.: On the benefits for model regularization of a Variational formulation of GTM. In: Proceedings of the IEEE International Joint Conference on Neural Networks (IJCNN 2008), Hong Kong (2008)

[55] Olier, I., Vellido, A.: A Variational Formulation for GTM Through Time. In: Proceedings of the IEEE International Joint Conference on Neural Networks (IJCNN 2008), Hong Kong (2008)

[56] P, P., H, P., K, S.: Neocortical glial cell numbers in human brain. Neuobiology of aging **29**, 1754–1762 (2008)

[57] P, T., RJ, G., Rothwell, J.: Arm function after stroke: Neurophysiological correlates and recovery mechanisms assessed by transcranial magnetic stimulation. Clinical Neurophysiology **117**, 1641–1659 (2006)

[58] Pallanti, S., Bernardi, S.: Neurobiology of repeated transcranial magnetic stimulation in the treatment of anxiety: A review. Int Clin Psycopharmacology **24**(4), 163–173 (2008)

[59] Rabiner, L.: A tutorial on Hidden Markov Models and selected applications in speech recognition. Proceedings of the IEEE **77**(2), 257–285 (1989)

[60] Rojo, N., Amengual, J., Rodriguez-Fornells, A.: Music-supported therapy induces plasticity in the sensorimotor cortex in chronic stroke patients: a case report using multimodal imaging (fmri-tms). Submitted for Publication in Neuror and Neural Rep (2010)

[61] Rossini, P., Rossini, L., Ferreri, F.: Transcranial magnetic stimulation: A review. IEEE Xplore **3**, 84–95 (2010)

[62] S, F. (ed.): Origins of neuroscience: a history of explorations into brain function. Oxford University Press, US (2001)

[63] Schneider, S., Schoenle, P., Altenmueller, E., Muente, T.: Using musica instruments to improve motor skill recovery following a stroke. J Neurol. **10**, 1339–1346 (2007)

[64] Svensén, M.: GTM: The Generative Topographic Mapping. Ph.D. thesis, Aston University (1998)

[65] Terao, Y., Ugawa, Y.: Basic mechanisms of tms. J Clin Neurophysiol. **19**(4), 322–243 (2002)

[66] Tergau, F., Wanchura, V., Canelo, M., Wischer, S., Wassermann, E., Ziemann, U.: Complete suppression of voluntary motor drive during the silent period after transcranial magnetic stimulation. Exp Brain Res **124**, 447–54 (1999)

[67] Tino, P., Nabney, I.: Hierarchical gtm: Constructing localized nonlinear projection manifolds in a principled way. IEEE Trans. Pattern Anal. Mach. Intell. **24**(5), 639–656 (2002)

[68] Traversa, R., Cicinelli, P., Olivery, M., Filippi, M., Rossini, P.: Neurophysiological follow-up of motor cortical output in stroke patients. Clin Neurophysiol **111**, 1695–1703 (2000)

[69] Triggs, W., Rosler, K., Truffert, A., Mayers, J.: Motor inhibition and excitation are independent effects of magnetic cortical stimulation. Annals of Neurology **32**, 345–351 (1992)

[70] Tseng, P., Hsu, T., Juan, C.: Posterior parietal cortex mediates encoding and maintenance processes in change blindness. Neuropsychologia **48**(4), 1063–1070 (2010)

[71] Twitchell, T.: The prognosis of motor recovery in hemiplegia. Bull tufts N Engl Med Cent **3**, 1860–1867 (1957)

[72] Vaynman, S., FGomez-Pinilla: License to run: exercise impacts functional plasticity in the intact and injured cns by using neurotrophins. Neurorehabil Neural Repair **19**, 283–295 (2005)

[73] Vellido, A., El-Deredy, W., Lisboa, P.J.G.: Selective smoothing of the Generative Topographic Mapping. IEEE T. Neural Networ. **14**(4), 847–852 (2003)

[74] Vellido, A., Lisboa, P.J.G., Vicente, D.: Robust analysis of MRS brain tumour data using t-GTM. Neurocomputing **69**(7-9), 754–768 (2006)

[75] Zhang, G., Patuwo, B., Hu, M.: Forecasting with artificial neural networks: The state of the art. Int. J. of Forecasting **14**(1), 35–62 (1998)