

UNIVERSIDAD NACIONAL DEL LITORAL



DOCTORADO EN INGENIERÍA

# **Decodificación de la actividad cerebral mediante regularización con penalizantes mixtos**

Victoria Peterson

**FICH**

FACULTAD DE INGENIERÍA Y CIENCIAS HÍDRICAS

**INTEC**

INSTITUTO DE DESARROLLO TECNOLÓGICO PARA LA INDUSTRIA QUÍMICA

**CIMEC**

CENTRO DE INVESTIGACIÓN DE MÉTODOS COMPUTACIONALES

**sinc(*i*)**

INSTITUTO DE INVESTIGACIÓN EN SEÑALES, SISTEMAS E INTELIGENCIA COMPUTACIONAL

Tesis de Doctorado **2018**



UNIVERSIDAD NACIONAL DEL LITORAL  
Facultad de Ingeniería y Ciencias Hídricas  
Instituto de Desarrollo Tecnológico para la Industria Química  
Centro de Investigación de Métodos Computacionales  
Instituto de Investigación en Señales, Sistemas e Inteligencia Computacional

# **DECODIFICACIÓN DE LA ACTIVIDAD CEREBRAL MEDIANTE REGULARIZACIÓN CON PENALIZANTES MIXTOS**

**Victoria Peterson**

Tesis remitida al Comité Académico del Doctorado  
como parte de los requisitos para la obtención  
del grado de  
DOCTOR EN INGENIERIA  
Mención Inteligencia Computacional, Señales y Sistemas  
de la  
UNIVERSIDAD NACIONAL DEL LITORAL

**2018**

Secretaría de Posgrado, Facultad de Ingeniería y Ciencias Hídricas, Ciudad Universitaria, Paraje "El Pozo",  
S3000, Santa Fe, Argentina





UNIVERSIDAD NACIONAL DEL LITORAL  
Facultad de Ingeniería y Ciencias Hídricas  
Instituto de Desarrollo Tecnológico para la Industria Química  
Centro de Investigación de Métodos Computacionales  
Instituto de Investigaciones en Señales, Sistemas e Inteligencia Computacional

## DECODIFICACIÓN DE LA ACTIVIDAD CEREBRAL MEDIANTE REGULARIZACIÓN CON PENALIZANTES MIXTOS

**Victoria Peterson**

### **Lugar de Trabajo:**

sinc(i)

Instituto de Investigación en Señales, Sistemas e  
Inteligencia Computacional  
Facultad de Ingeniería y Ciencias Hídricas  
Universidad Nacional del Litoral

### **Director:**

Dr. Rubén Daniel Spies

IMAL-FIQ-UNL-CONICET

### **Co-director:**

Dr. Hugo Leonardo Rufiner

sinc(i)-FICH-UNL-CONICET

### **Jurado Evaluador:**

Dr. Diego Tomassi

FIQ-UNL-CONICET

Dr. José Biurrun Manresa

IBB-FIUNER-CONICET

Dr. Pablo Diez

UNSJ/CONICET

Dr. Carlos Muravchik

LECI-CONICET/UNLP

**2018**



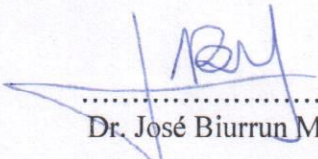



UNIVERSIDAD NACIONAL DEL LITORAL  
Facultad de Ingeniería y Ciencias Hídricas

Santa Fe, 2 de Noviembre de 2018.


Como miembros del Jurado Evaluador de la Tesis de Doctorado en Ingeniería titulada *"Decodificación de la actividad cerebral mediante regularización con penalizantes mixtos"*, desarrollada por la Bioing. Victoria PETERSON, en el marco de la Mención "Inteligencia Computacional, Señales y Sistemas", certificamos que hemos evaluado la Tesis y recomendamos que sea aceptada como parte de los requisitos para la obtención del título de Doctor en Ingeniería.

La aprobación final de esta disertación estará condicionada a la presentación de dos copias encuadradas de la versión final de la Tesis ante el Comité Académico del Doctorado en Ingeniería.

  
.....  
Dr. José Biurrún Manresa

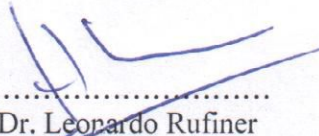
  
.....  
Dr. Carlos Muravchik

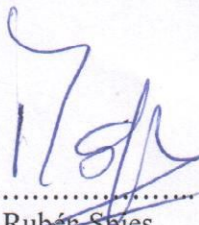
.....  
Dr. Pablo Diez (\*)

  
.....  
Dr. Diego Tomassi

Santa Fe, 2 de Noviembre de 2018

Certifico haber leído la Tesis, preparada bajo mi dirección en el marco de la Mención "Inteligencia Computacional, Señales y Sistemas" y recomiendo que sea aceptada como parte de los requisitos para la obtención del título de Doctor en Ingeniería.

  
.....  
Dr. Leonardo Rufiner  
Codirector de Tesis

  
.....  
Dr. Rubén Spies  
Director de Tesis

(\*) El Dr. Pablo Diez participó por videoconferencia

  
Dr. JOSÉ LUIS MACOR  
SECRETARIO DE POSGRADO  
Facultad de Ingeniería y Cs. Hídricas

Universidad Nacional del  
Litoral  
Facultad de Ingeniería y  
Ciencias Hídricas

Secretaría de Posgrado

Ciudad Universitaria  
C.C. 217  
Ruta Nacional Nº 168 - Km. 472,4  
(3000) Santa Fe  
Tel: (54) (0342) 4575 229  
Fax: (54) (0342) 4575 224  
E-mail: posgrado@fich.unl.edu.ar



## **DECLARACIÓN DEL AUTOR**

Esta tesis ha sido remitida como parte de los requisitos para la obtención del grado académico de Doctor en Ingeniería - Mención Inteligencia Computacional, Señales y Sistemas ante la Universidad Nacional del Litoral y ha sido depositada en la Biblioteca de la Facultad de Ingeniería y Ciencias Hídricas para que esté a disposición de sus lectores bajo las condiciones estipuladas por el Reglamento de la mencionada Biblioteca.

Citaciones breves de esta tesis son permitidas sin la necesidad de un permiso especial, en la suposición de que la fuente sea correctamente citada. Solicitudes de permiso para la citación extendida o para la reproducción parcial o total de este manuscrito serían concebidos por el portador legal del derecho de propiedad intelectual de la obra.





*-Florecer exige pasar por todas las estaciones-*  
PARA AQUELLOS QUE SIEMBRAN, RIEGAN Y FLORECEN.



# Agradecimientos

“No alcanzan las palabras cuando lo que hay que decir desborda el alma”, decía Julio Cortázar. Y con el alma desbordada, de resiliencia evolucionada, de alegría inmensa por esos micro logros y por cada caída recuperada, es que intento agradecer en un orden que espero que no altere el producto.

En primer lugar, y sin un orden realmente establecido, quiero agradecer a mis Directores, que confiaron en mí y me brindaron las herramientas para recorrer el camino. En especial a Rubén, que allá por 2013 se animó a dirigir a una Ingeniera, que no conocía pero que se ve que el destino algo tenía preparado...gracias infinitas por el apoyo constante, los mates (fríos) y las largas tardes de pestañas quemadas arriba de un Paper.

En segundo lugar, y sin un orden realmente establecido, agradecer a mí novio Emi Ravera, mi fiel compañero que soportó todos los estados que una becaria doctoral puede desarrollar a lo largo de una tesis. A mis amigos, la familia que se elige, los que uno va sembrando y recolectando en el camino, por cada chiste nerd que me dejaron hacer/reír. A mi familia, la de sangre, por dejarme crecer, y bajarme a tierra los domingos de asados.

En tercer lugar, y sin un orden realmente establecido, agradecer a Roger Gassert, Olivier Lamercy, Dominik Wyser y todo el equipo del RELab-ETHZ, por todo lo que recibí en esos seis meses junto a ellos, desde lo académico, lo humano y hasta los ricos chocolates Suizos. A las personas que allí conocí y que hicieron que todo sea más fácil. También a la Institución Zeno Karl Schindler Foundation (ZKS Foundation, Suiza) por confiar en mí y otorgarme el subsidio suficiente para que ese viaje sea posible.

Y hablando de instituciones, en último lugar, y sin un orden realmente establecido, a cada Institución que me dio su apoyo, desde el mismo CONICET, invirtiendo en mí, hasta el sinc(i) y el IMAL, con las personas que lo forman y constituyen, especialmente a aquellas personas que la dirigen, quienes me permitieron tener una pata en cada mundo.

La felicidad no es la meta cumplida, es el camino recorrido. Esta condición necesaria, pero no suficiente, para poder simplemente lograr una única cosa: investigar.



# Acrónimos

En el presente documento, para la mayoría de las abreviaturas, utilizaremos los acrónimos correspondientes a las expresiones en idioma inglés, por ser la forma más difundida y utilizada en la bibliografía.

---

ACV	Accidente cerebro-vascular.
AUC	Área bajo la curva ROC.
BCDF	Función distribución binomial acumulativa.
BCI	Interfaces cerebro-computadora.
CNS	Sistema nervioso central.
CSP	Patrones espaciales comunes.
DFBCSP	Patrones espaciales comunes y banco de filtros discriminativos.
EEG	Electroencefalografía.
EMG	Electromiografía.
ENET	Red elástica ( <i>elastic-net</i> ).
ERD	Evento relacionado a la desincronización.
ERP	Potencial relacionado a eventos.
ERS	Evento relacionado a la sincronización.
FBCSP	Patrones espaciales comunes y banco de filtros.
FC	Criterio de Fisher.
FLDA	Análisis discriminante lineal de Fisher.
LASSO	Operador de mínima reducción y selección absoluta.
LDA	Análisis discriminante lineal.
LSR	Regresión de mínimos cuadrados.
MI	Imaginería motora.
PNS	Sistema nervioso periférico.
PFBCSP	Patrones espaciales comunes y bandas frecuenciales penalizadas.
PTFBCSP	Patrones espaciales comunes y bandas temporo-frecuenciales penalizadas.

SFBCSP	Patrones espaciales comunes y banco de filtros ralos.
SDA	Análisis discriminante ralo.
SKLDA	Análisis discriminante lineal de contracción.
SMR	Ritmos sensoriomotores.
SNR	Relación señal a ruido.
SWLDA	Análisis discriminante lineal paso por paso.

---

# Notación

---

$\mathbf{X}$	matriz de datos de dimensión $n \times p$ .
$\mathbf{x}_i$	$i$ -ésima realización de dimensión $p$ .
$K$	cantidad de clases.
$\mathbf{y}$	vector de etiquetas.
$\mathbf{Y}$	matriz de etiquetas binarias, $\mathbf{Y}_i \in \{(0, 1), (1, 0)\}$ .
$\Sigma$	matriz de covarianza.
$\mu$	media muestral.
$I_k$	conjunto de índices de las observaciones pertenecientes a la $k$ -ésima clase.
$ I_k  = n_k$	cantidad de observaciones pertenecientes a la $k$ -ésima clase.
$\beta$	vector solución.
$\theta$	vector de puntajes.
$\lambda$	parámetro de regularización.
$f_k$	función distribución de probabilidad de la $k$ -ésima clase.
$\mathbf{D}$	matriz de anisotropía.

---





# Índice general

<b>1. Introducción</b>	<b>1</b>
1.1. Motivación . . . . .	1
1.2. Objetivos . . . . .	4
1.2.1. Objetivos Particulares . . . . .	4
<b>2. Interfaces cerebro-computadora</b>	<b>5</b>
2.1. Bases neurofisiológicas . . . . .	5
2.1.1. Sistema nervioso central . . . . .	6
2.1.2. Sistema nervioso periférico . . . . .	10
2.2. Una nueva forma de comunicación . . . . .	10
2.2.1. Midiendo la actividad cerebral . . . . .	11
2.2.2. El lenguaje del cerebro . . . . .	13
2.3. Detectando la actividad cerebral . . . . .	17
2.3.1. Pre-procesamiento . . . . .	18
2.3.2. Extracción de características . . . . .	19
2.3.3. Selección de características . . . . .	20
2.3.4. Clasificación . . . . .	20
2.4. Estado del arte y desafíos actuales . . . . .	21
<b>3. Datos reales: registros de EEG en BCI</b>	<b>25</b>
3.1. Base de datos de EEG bajo el paradigma P300 . . . . .	25
3.1.1. P300-1: Base de datos ERP-EEG del LINI-UAM, México . . . . .	26
3.1.2. P300-2: Base de datos II de la Competencia III de BCI . . . . .	26
3.2. Base de datos de EEG bajo el paradigma de MI . . . . .	27
3.2.1. MI-1: Base de datos propia . . . . .	27
3.2.2. MI-2: Base de datos IVa de la Competencia III de BCI . . . . .	28
3.2.3. MI-3: Base de datos IIb de la Competencia IV de BCI . . . . .	28

<b>4. Discriminación rala y regularización con penalizantes mixtos</b>	<b>31</b>
4.1. Análisis discriminante lineal: su concepción y evolución	31
4.2. Análisis discriminante lineal ralo generalizado	35
4.3. Experimentos y resultados	39
4.3.1. KLD y matrices de anisotropía	39
4.3.2. Desempeño de GSDA vs. SDA	42
4.4. Discusiones	45
<b>5. Escenario de entrenamiento pequeño: un enfoque para optimizar los tiempos de calibración en BCIs basadas en ERPs</b>	<b>47</b>
5.1. La importancia de disminuir los tiempos de calibración	47
5.1.1. Experimentos y resultados	49
5.2. Medidas de discrepancia: un análisis de su impacto en la detección de ERPs	51
5.2.1. Experimentos y resultados	52
5.3. Discusiones	54
<b>6. Un método integrador para la detección de la intención de movimiento</b>	<b>59</b>
6.1. Filtrado espacial y banco de filtros	59
6.2. Patrones espaciales comunes y bandas temporo-frecuenciales penalizadas	61
6.3. Experimentos y resultados	64
6.3.1. Análisis de los datos	65
6.3.2. Resultados de clasificación	66
6.3.3. Mapas topográficos de los patrones temporo-frecuenciales espaciales	69
6.4. Discusiones	72
<b>7. Hacia la aplicación final: evaluación en tiempo real de una BCI basada en MI</b>	<b>75</b>
7.1. Introducción	75
7.2. Descripción del experimento	76
7.2.1. Protocolo de estimulación	77
7.2.2. Software y hardware utilizados	79
7.2.3. Adquisición de datos	80
7.3. Resultados	82
7.3.1. Análisis de la señal de EMG	82
7.3.2. Resultados de clasificación en tiempo real	83
7.4. Discusiones	85

---

<b>8. Conclusiones y trabajos futuros</b>	<b>87</b>
<b>Anexos</b>	<b>91</b>
<b>A. Detalles sobre FLDA y CSP</b>	<b>93</b>
A.1. FLDA como un problema de mínimos cuadrados . . . . .	93
A.2. Método CSP . . . . .	93
<b>B. Ad-hoc Gaussian Dictionaries for Sparse Representation of Evoked Re- lated Potentials</b>	<b>95</b>
<b>C. <math>\ell_1</math>-norm Regularization for Sparse Representation and P300 Wave De- tection in Brain-Computer Interfaces</b>	<b>109</b>



# Índice de figuras

1.1. Diagrama de bloques de una BCI genérica. . . . .	3
2.1. Vista general de la corteza cerebral, lóbulos y cisuras. . . . .	7
2.2. Esquema representativo de la localización de las áreas cerebrales más relevantes en la corteza cerebral. . . . .	7
2.3. Mapa de la corteza cerebral de las zonas asociadas al sentido y control de ciertas partes del cuerpo, representado mediante el homúnculo de Penfield. . . . .	8
2.4. Vista lateral y superior del sistema internacional 10-20 de localización de electrodos junto con referencias de distancia y nombres de los electrodos más relevantes. . . . .	13
2.5. Señal ERP con sus ondas componente. Notar que, por convención, se adopta polaridad positiva hacia abajo. . . . .	14
2.6. Patrones ERD/ERS observables antes, durante y después de la ejecución del movimiento, analizado en diferentes ritmos frecuenciales (10-12 Hz, 14-18 Hz y 36-40 Hz). La señal de EMG marca el comienzo del movimiento. Electrodo de análisis marcado en negro. . . . .	16
4.1. Divergencia-J (KLD) y matrices de anisotropía ( $\mathbf{D}_1$ y $\mathbf{D}_2$ ) para diferentes sujetos de las bases de datos P300-1 y P300-2. . . . .	41
4.2. Valores del área bajo la curva ROC (AUC) alcanzados en los datos de testeo al utilizar SDA y GSDA para la base de datos P300-1 evaluada mediante validación cruzada de 3 particiones. Para cada sujeto, las barras y las líneas negras representan el promedio y el desvío estándar, respectivamente, de los valores de AUC en las tres particiones. La última columna corresponde al promedio del desempeño de todos los sujetos. Aquí “**” indica que GSDA>SDA con $p$ -valor< 0,05. . . . .	43

4.3. Cantidad de iteraciones requeridas por SDA y GSDA para encontrar el vector solución $\hat{\beta}$ para la base de datos P300-1 (a) y P300-2 (b).	45
5.1. Promedio de AUC para los datos de testeo alcanzados por LDA, SWLDA, SKLDA, FC+LDA y GSDA, respectivamente, en cada uno de los escenarios de entrenamiento pequeño simulado para las bases de datos P300-1 y P300-2. La barra de error indica la desviación estándar de AUC a lo largo de las 100 repeticiones.	50
5.2. Grandes promedios y gráficas discriminativas para un sujeto de la base de datos P300-1. De izquierda a derecha y de arriba a abajo se presenta en el plano tiempo-canal los mapas discriminativos calculados mediante: promediación coherente para la clase con P300, promediación coherente para clase sin P300, $d_J$ , $d_{JS_{\pi_1, \pi_2}}$ , $d_V$ , $d_E$ , $d_F$ y $d_r$ . a) Escenario-1 b) Escenario-6	56
5.3. Grandes promedios y gráficas discriminativas para un sujeto de la base de datos P300-2. De izquierda a derecha y de arriba a abajo se presenta en el plano tiempo-canal los mapas discriminativos calculados mediante: promediación coherente para la clase con P300, promediación coherente para clase sin P300, $d_J$ , $d_{JS_{\pi_1, \pi_2}}$ , $d_V$ , $d_E$ , $d_F$ y $d_r$ . a) Escenario-1 b) Escenario-6	57
5.4. Resultados de clasificación promedio en términos de AUC en cada escenario de entrenamiento simulado al utilizar las diferentes medidas de discrepancia consideradas dentro de GSDA.	58
6.1. Representación esquemática del método PTFBCSP propuesto. La señal de EEG multicanal es analizada en $T$ ventanas temporales. Luego, cada $t$ -ésimo segmento temporal es dividido en $F$ bandas frecuenciales. El método CSP es aplicado a cada segmento temporo-frecuencial. Finalmente, la selección de las características CSP temporo-frecuenciales óptimas es automáticamente realizada junto con la clasificación mediante GSDA.	62
6.2. Resultados de clasificación promedio sobre las $10 \times 10$ validaciones cruzadas obtenidas mediante CSP, SFBCSP, PFBCSP y PTFBCSP para la base de datos MI-1. Aquí “*” denota diferencia estadística significativa entre los métodos.	67
6.3. Resultados de clasificación promedio sobre las $10 \times 10$ validaciones cruzadas obtenidas mediante CSP, SFBCSP, PFBCSP y PTFBCSP para la base de datos MI-2. Aquí “*” denota diferencia estadística significativa entre los métodos.	68

6.4. Resultados de clasificación promedio sobre las $10 \times 10$ validaciones cruzadas obtenidas mediante CSP, SFBCSP, PFBCSP y PTFBCSP para la base de datos MI-3. Aquí “*” denota diferencia estadística significativa entre los métodos. . . . .	69
6.5. Arriba: características seleccionadas por cada método de ventana única analizado. Abajo: mapas topográficos de filtro y el patrón espacial correspondiente a la característica más relevante (resaltada en rojo) para un sujeto de la MI-1 con dominancia izquierda. a) SFBCSP, b) PFBCSP. . . .	70
6.6. Arriba: características espectro-temporales espaciales seleccionadas por el método propuesto de banco de filtros y múltiples ventanas temporales en cada segmento temporal $t = 1, 2, \dots, 5$ (uno por fila). Abajo: mapa topográfico del filtro y del patrón espacial asociado a la característica más relevante (resaltada en rojo) para un sujeto de la MI-1 con dominancia izquierda. . . . .	71
7.1. Representación esquemática de una época del experimento con la referencia de los tiempos del protocolo, en segundos. . . . .	79
7.2. Localización de los 64 electrodos utilizados para la toma de datos. Nótese que CPz ha sido utilizado como electrodo de referencia y AFz como tierra. . . . .	80
7.3. Descripción esquemática del experimento. Electrodo, equipo de adquisición y softwares utilizados para el registro de cada señal biomédica considerada. . . . .	81
7.4. Montaje del experimento para un sujeto con dominancia izquierda. Se muestran los sistemas de adquisición de EEG y EMG, como así también el protocolo de estimulación con la señal visual para una época de MI . . . .	82
7.5. Resultados de clasificación en cada ronda de evaluación para los 10 sujetos sanos. Se muestran las dos sesiones por sujeto. Las líneas discontinuas demarcan la separación entre sujeto y sujeto. La línea roja denota el mínimo valor de clasificación para controlar una BCI. . . . .	84
7.6. Resultados de clasificación en cada ronda de evaluación para el paciente con ACV. La línea roja denota el mínimo valor de clasificación para controlar una BCI. . . . .	85





## Índice de tablas

4.1. Promedio y desvío estándar de los valores del área bajo la curva ROC (AUC) alcanzados por SDA y GSDA en los datos de testeo para la base de datos P300-2 evaluada mediante validación cruzada de 3 particiones. . . . .	42
6.1. Descripción general de los valores promedio de clasificación y sus correspondientes desvíos estándar entre paréntesis. Valores expresados en porcentajes. Los mejores resultados de clasificación están resaltados en negrita. Aquí “*” denota que $PTFBCSP > SFBCSP$ con un $p$ -valor $< 0,01$ . . . . .	68
7.1. Criterios de inclusión y exclusión para sujetos sanos (arriba) y pacientes de ACV (abajo). . . . .	78



# Decodificación de la actividad cerebral mediante regularización con penalizantes mixtos

Victoria Peterson

Director de Tesis: Rubén Daniel Spies

Co-Director de Tesis: Hugo Leonardo Rufiner

## Resumen

Una interfaz cerebro-computadora (BCI, por sus siglas en inglés) es un sistema que provee una alternativa forma de comunicación entre el cerebro de una persona y el mundo exterior. Una forma eficiente y no invasiva de medir la actividad cerebral es mediante electroencefalografía (EEG) de superficie. Si el objetivo es deletrear palabras, suelen utilizarse BCIs basadas en los denominados potenciales relacionados a eventos (ERPs, por sus siglas en inglés). Para fines de rehabilitación, la mayoría de las BCIs se basan en el paradigma de imaginación motora (MI, por sus siglas en inglés), es decir, en la simulación mental de la realización de ciertos movimientos. En ambos paradigmas, la detección de la intención del usuario, inmersa en el EEG, puede abordarse de manera simplificada como un problema de reconocimiento de patrones binario.

El análisis discriminante lineal (LDA, por sus siglas en inglés) es un método de clasificación muy conocido y utilizado en el contexto de aprendizaje supervisado. Si bien LDA generalmente resulta en buenos desempeños de clasificación manteniendo la solución sencilla, el método falla cuando el número de muestras es relativamente grande en relación a la cantidad de observaciones. Varios autores, tanto de la comunidad de BCI, así como de Estadística aplicada, han propuesto diferentes versiones regularizadas de LDA, mostrando siempre las ventajas del uso de tales técnicas.

En esta tesis se ha desarrollado una versión penalizada y regularizada de LDA, a la que denominamos análisis discriminante raro generalizado (GSDA, por sus siglas en inglés), para clasificación binaria. Este método selecciona automáticamente las características relevantes de la señal de interés junto con la clasificación, teniendo en cuenta información a-priori de discrepancia entre clases. Su algoritmo de implementación está diseñado para seleccionar automáticamente los parámetros de regularización. Los experimentos numéricos con bases de datos de ambos paradigmas (ERP-EEG y MI-EEG) muestran que la utilización de GSDA supera el desempeño alcanzado por los métodos del estado del arte para clasificación tanto de ERP como de MI. Asimismo, como cierre de esta tesis, se presenta un estudio de la factibilidad de un método basado en GSDA para la detección de la intención del movimiento en tiempo real.



# Brain signal decoding based on mixed-norm regularization

Victoria Peterson

Thesis Director: Rubén Daniel Spies

Thesis Co-Director: Hugo Leonardo Rufiner

## Abstract

A brain computer interface (BCI) is a system which provides an alternative way of communication between the mind of a person and the outside world by using only measured brain activity. An efficient and non-invasive way of establishing the communication is based on electroencephalography (EEG) and event-related potentials (ERPs). An ERP is an endogenous potential which results as a consequence of an external and relevant stimulus. For rehabilitation purposes most of the BCIs are based on motor imagery (MI), i.e. the mental simulation of movements. For the pattern recognition point of view both paradigms represent an extremely hard and challenging binary classification problem.

The linear discriminant analysis (LDA) criterion is a well-known and widely used dimensionality reduction tool in the context of supervised classification. Although LDA generally results in good classification performances while keeping the solution simple, it fails when the number of samples is large relative to the number of observations. Several authors, both from the BCI and the statistical research communities, have proposed different regularized versions of LDA, showing always the advantages of such tools.

In this thesis we present the generalized sparse discriminant analysis (GSDA) framework, for binary classification. This method automatically performs discriminative feature selection and classification by taking into account a-priori class discrepancy information. The GSDA method is designed to automatically select the optimal regularization parameters. Numerical experiments with both ERP-EEG and MI-EEG datasets are presented, showing that overall GSDA performance outperforms most state-of-the-art ERP and MI classification algorithms, for single-trial EEG classification. In addition, at the end of this thesis, a feasibility study of a proposed method based on GSDA for MI detection in real-time scenarios is presented.



# Capítulo 1

## Introducción

“No podés hablar de etapas sin presuponer una meta.  
— Llamáale hipótesis de trabajo, cualquier cosa así”

---

Julio Cortázar, Rayuela, Capítulo 99

### 1.1. Motivación

Enfermedades como la esclerosis lateral amiotrófica, los accidentes cerebro vasculares, la esclerosis múltiple, el traumatismo medular o cerebral, la parálisis cerebral, o la distrofia muscular, por nombrar algunas, alteran las vías neuronales que controlan los músculos. Este tipo de enfermedades disminuyen o anulan la capacidad de los canales neuromusculares por los cuales el cerebro se comunica normalmente y/o controla su entorno. Independientemente de la causa, la pérdida del control motor puede tener diversos grados de severidad, desde la paresia de un grupo de músculos hasta la pérdida de control total de toda la musculatura incluyendo el control ocular, siendo este último caso conocido como “síndrome de enclaustramiento”. Generalmente, el individuo afectado pierde su capacidad de comunicación o interacción con el medio a pesar de tener sus capacidades cognitivas inalteradas.

Actualmente muchas de estas lesiones no pueden repararse totalmente mediante técnicas quirúrgicas o terapéuticas, produciendo que el paciente quedé con disfunciones motoras permanentes, y en casos severos, sin capacidad de comunicación alguna. Por lo tanto, surge la necesidad de establecer un nuevo canal de comunicación entre el individuo y su medio, que no dependa del accionar de las vías nerviosas o musculares normales. Las denominadas interfaces cerebro-computadora (BCI, por sus siglas en inglés) son dispositivos



que permiten establecer una vía de comunicación directa y alternativa, entre el cerebro de la persona y el mundo exterior, siendo actualmente dispositivos de comunicación y herramientas de rehabilitación prometedores.

Para aquellos individuos con poca o nula actividad motora remanente, las BCI basadas en señales de electroencefalografía (EEG) permiten que la realización de una acción o el comando de un dispositivo externo dependa solamente de la intención de los usuarios. Existen diferentes paradigmas para transformar la información contenida en la señal de EEG a comandos de control o comunicación. Dos de los paradigmas más estudiados son el basado en ritmos sensoriomotores y el basado en potenciales evocados relacionados a eventos.

Independientemente del tipo de paradigma utilizado, toda BCI puede describirse mediante un diagrama en bloques. La Figura [1.1](#) representa el esquema de una BCI genérica, en el cuál se destaca el bloque de adquisición, encargado de medir la actividad cerebral (señal de entrada), el bloque de decodificación, en el cual la señal de entrada se procesa y traduce a señales de control (señal de salida), y finalmente, un bloque con la aplicación de la BCI, en el cuál los comandos generan ciertas acciones en el dispositivo a controlar. Una intercomunicación entre el dispositivo a controlar y el usuario de la BCI, se genera gracias a los estímulos y retroalimentación que recibe el usuario.

En los últimos años la comunidad de BCI ha puesto mucho esfuerzo en el diseño de algoritmos de traducción de la señal de EEG en comandos de control. Diversas herramientas de procesamiento de señales, reconocimiento de patrones y aprendizaje maquina, entre otras, han sido construidas y utilizadas para poder brindar métodos estables, confiables, robustos al ruido, adaptables a las variaciones del sujeto, veloces en la traducción de la señal e implementables en tiempo real. A pesar de los avances obtenidos, aún queda mucho espacio para mejorar el desempeño de los métodos actuales. A lo largo de esta tesis, si bien se abordaron todos los bloques de una BCI antes mencionados, nos centramos principalmente en la construcción de algoritmos de traducción de la señal de EEG basados en regularización con penalizantes mixtos para decodificar correctamente la intención del usuario.

En una primera instancia de esta investigación, enfocados en resolver el problema de detección de ERPs en la señal de EEG, se evaluó el uso de técnicas de clasificación basadas en representación rara y diccionarios ad-hoc construidos con funciones Gaussianas bidimensionales. Si bien los resultados encontrados no superaban el estado del arte, constituyó de un trabajo preliminar que probablemente pueda mejorarse mediante una profundización y mejor caracterización de la representación de la señal de interés me-

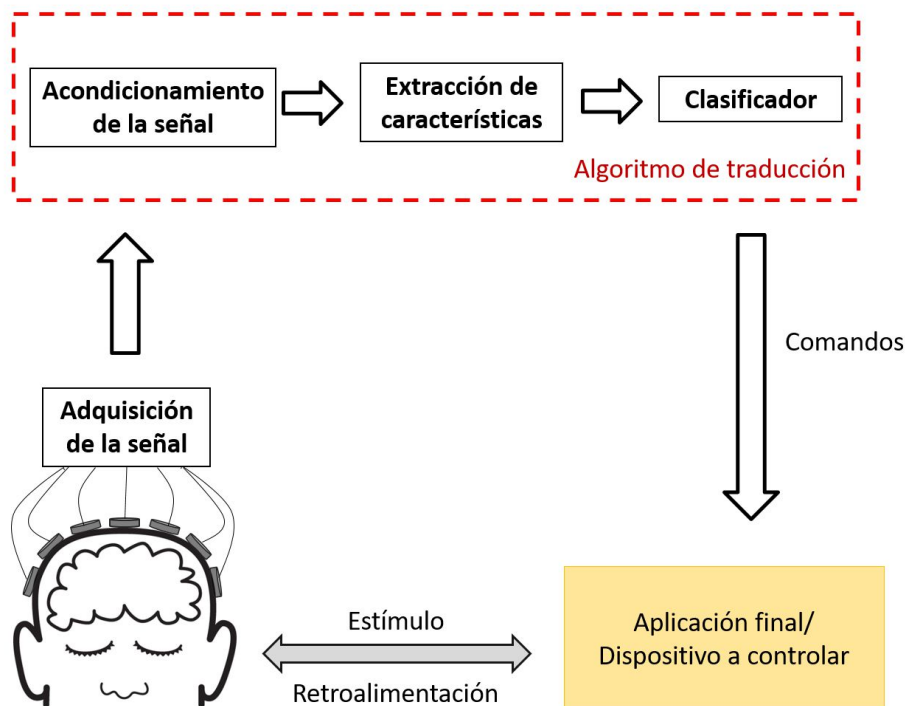


Figura 1.1: Diagrama de bloques de una BCI genérica.

dante ese u otro tipo de diccionario más complejo. El método propuesto junto, con los resultados experimentales, fue publicado en congresos de alcance internacional y nacional (Peterson et al., 2015a,b). Si bien en el cuerpo principal de esta tesis no se presentan detalles sobre dicha metodología, en los Anexos B y C se adjuntan, respectivamente, dichas publicaciones en su formato e idioma originales.

Este trabajo de tesis está organizado la siguiente manera. A continuación se detallan los objetivos perseguidos a lo largo de la misma. En el Capítulo 2 se realiza una breve introducción a las interfaces cerebro-computadora, detallando sus desafíos actuales. El Capítulo 3 describen los materiales (base de datos) utilizados en los experimentos numéricos realizados a lo largo de este trabajo. En el Capítulo 4 se presenta el principal método de clasificación desarrollado. Los capítulos 5 y 6 presentan los experimentos y resultados de la utilización del clasificador propuesto en dos paradigmas de BCI bien conocidos. Una breve descripción de la evaluación en tiempo real del clasificador es realizada en el Capítulo 7. Finalmente, conclusiones y trabajos futuros son presentados en el Capítulo 8.

## 1.2. Objetivos

Como objetivo general nos planteamos desarrollar algoritmos avanzados de procesamiento y reconocimiento de señales basados en métodos de regularización con penalizantes mixtos, que sean aplicables a señales derivadas de la actividad cerebral en el contexto de las BCI, cuyo tratamiento con los métodos actuales resulta difícil, lento, insuficiente y/o no permite extraer conclusiones satisfactorias.

### 1.2.1. Objetivos Particulares

- Adquirir los conocimientos necesarios para el análisis y procesamiento avanzado de señales y reconocimiento de patrones.
- Desarrollar nuevos algoritmos de clasificación que permitan extraer información relevante de la señal de EEG.
- Introducir información a-priori sobre la discrepancia entre clases en un modelo discriminativo integrado.
- Estudiar el impacto del uso de diferentes medidas de discrepancia entre clases en el desempeño final del clasificador.
- Construir un clasificador robusto, rápido y confiable que permita disminuir los tiempos de calibración en los sistemas BCIs basados en ERPs.
- Construir un algoritmo sujeto-específico de detección de la intención del movimiento.
- Construir una base de datos propia de registros de EEG utilizando equipos confiables y transportables.
- Realizar una evaluación del éxito de las diferentes técnicas de detección y clasificación propuestas utilizando bases de datos de libre acceso como así también una base de datos propia.
- Realizar un evaluación en tiempo real del método de detección propuesto.

## Capítulo 2

# Interfaces cerebro-computadora

“Empieza por el principio,” dijo el rey con gravedad,  
“y sigue hasta llegar al final: entonces para.”

---

Lewis Carroll, Alicia en el País de las maravillas

Por miles de años nuestro cerebro ha evolucionado para controlar eficientemente un gran y complejo dispositivo biológico: el cuerpo humano. Actualmente, la versatilidad de nuestro sistema nervioso ha llevado a que seamos capaces de extender esta capacidad de control a dispositivos externos muy diferentes a nuestro propio cuerpo. Las interfaces cerebro-computadora (BCIs, por sus siglas en inglés), tema central de este capítulo, constituyen un área interdisciplinaria que persigue este objetivo aprovechándose de avances recientes en neurociencia, procesamiento de señales, aprendizaje maquina, tecnologías de la información, entre otros. En este capítulo se presentan brevemente las bases neurofisiológicas, los tipos de BCI, su arquitectura general, así como sus desafíos y problemáticas actuales.

### 2.1. Bases neurofisiológicas<sup>1</sup>

Las neuronas conforman la unidad básica del sistema nervioso. Este tipo de células especializadas están compuestas por un cuerpo celular (llamado soma), las dendritas y el axón. La comunicación, química o eléctrica, entre neuronas se denomina sinapsis, proceso

---

<sup>1</sup>La información contenida en esta sección es un breve resumen que describe los conceptos neurofisiológicos básicos detallados en la bibliografía de referencia para una mejor comprensión de la relevancia y funcionalidad de los sistemas BCIs. Para la escritura, en particular nos hemos basado en (Clerc et al. 2016; Lataret y Liard, 2004; Rao, 2013).

en el cual la información (inhibitoria o excitatoria) viaja en forma de potencial de acción (también llamado impulso nervioso) con rapidez y precisión, alcanzando largas distancias. Billones de neuronas interconectadas entre sí conforman el tejido nervioso, una densa red biológicamente distinguible en sustancia gris y sustancia blanca, correspondientes a las regiones que contienen los cuerpos celulares y las fibras nerviosas mielinizadas (cubiertas de mielina) de las neuronas, respectivamente.

El sistema nervioso humano puede ser dividido esencialmente en dos grandes sistemas: el sistema nervioso central (CNS, por sus siglas en inglés) y el sistema nervioso periférico (PNS, por sus siglas en inglés). El PNS está constituido por el *sistema nervioso somático* (nervios craneales y espinales que envían información sensorial al CNS) y el *sistema nervioso autónomo* (rama simpática y parasimpática que regula las funciones viscerales como la respiración y el bombeo sanguíneo). El CNS incluye al encéfalo, localizado en la cavidad craneana, y la médula espinal, localizada en el canal vertebral.

### 2.1.1. Sistema nervioso central

El cerebro se compone de dos hemisferios, derecho e izquierdo, que están conectados esencialmente por el haz de fibras nerviosas (sustancia blanca) denominado cuerpo caloso. La corteza cerebral, tejido nervioso que cubre la superficie de los hemisferios cerebrales, conforma una delgada capa de sustancia gris fuertemente circunvolucionada. En ella es posible distinguir el lóbulo frontal, el lóbulo parietal, el lóbulo occipital y el lóbulo temporal. Estos lóbulos están separados por cisuras, cuyas dos más importantes son el surco central (o cisura de Rolando) y el surco lateral (o cisura de Silvio). En la Figura [2.1](#) se representan esquemáticamente dichos lóbulos y cisuras.

Cada lóbulo presenta diferentes características y se lo asocia a diferentes áreas y funciones. La Figura [2.2](#) muestra los lóbulos con sus áreas correspondientes. A continuación se describe brevemente cada uno de ellos.

#### Lóbulo frontal

Se puede dividir en diferentes regiones citoarquitectónicas y funcionales: la corteza motora y prefrontal.

- **Corteza motora:** la corteza motora compone de área motora primaria y áreas premotoras, el campo ocular frontal y el área de Broca. El **área motora primaria** (M1) se encuentra delante de la circunvolución precentral. Esta zona es la que

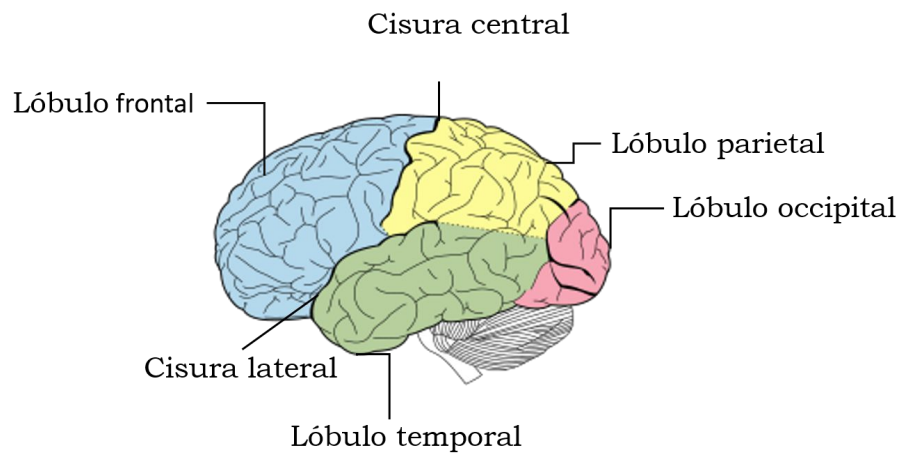


Figura 2.1: Vista general de la corteza cerebral, lóbulos y cisuras.

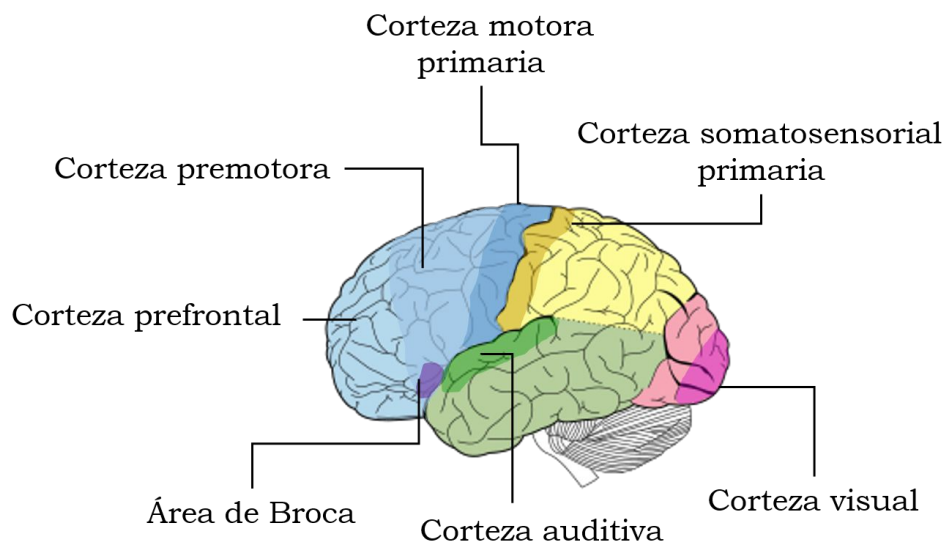


Figura 2.2: Esquema representativo de la localización de las áreas cerebrales más relevantes en la corteza cerebral.

proporciona la representación y ejecución de las diferentes zonas del cuerpo, representadas por el homúnculo de Penfield (Penfield y Boldrey, 1937) (ver Figura 2.3). Una lesión en el córtex motor primario causaría una parálisis contralateral a la lesión (Clerc et al., 2016).

El **área premotora** controla principalmente la musculatura del tronco y la musculatura proximal de las extremidades y está involucrada en la preparación, iniciación

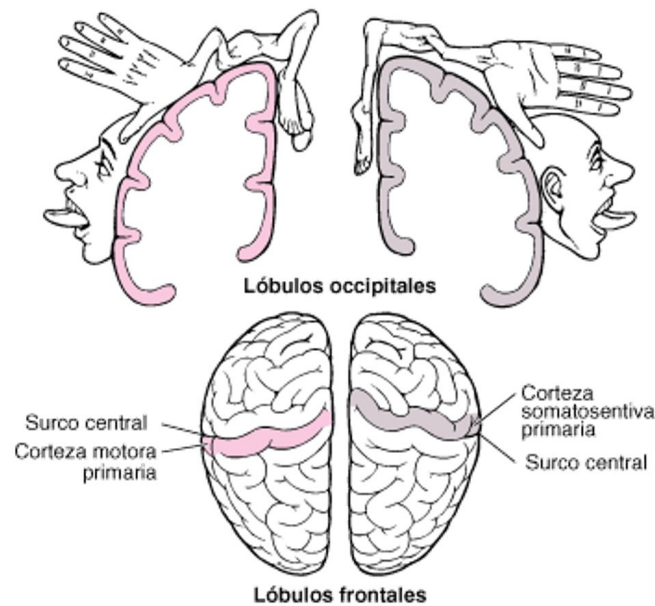


Figura 2.3: Mapa de la corteza cerebral de las zonas asociadas al sentido y control de ciertas partes del cuerpo, representado mediante el homúnculo de Penfield.

y monitoreo de movimientos complejos. Una lesión en el área premotora provocaría acinesia (pérdida) del lado contralateral a la lesión, reducción del habla espontánea y reducción de la actividad motora espontánea contralateral. Puede provocar movimientos lentos y no coordinados, dificultad de programar movimientos alternos, rítmicos y secuenciales.

El **campo ocular frontal** recibe proyecciones de las áreas visuales y auditivas. Controla los movimientos oculares voluntarios. Esto le permite realizar una correcta focalización y atención visual. Una lesión en el campo ocular frontal causaría una alteración en el sistema visomotor provocando un enlentecimiento en el rastreo visual, una localización incorrecta de los sonidos y alteración de la atención visual.

El **área de Broca** se encarga del plan de secuenciación motora del habla. Una lesión provocaría una grave pérdida de la fluidez verbal.

- **Corteza prefrontal:** la corteza prefrontal juega un rol importante en la elaboración de pensamientos y en el desarrollo de la personalidad. Es la última región desarrollada y evolucionada del cerebro. Esta corteza recibe numerosas aferencias procedentes de todo el cerebro, y es por esto que juega un papel esencial en las emociones, la con-

ducta, la motivación y ejecución de ideas, así como en la toma de decisiones. Daños en esta zona cerebral pueden producir desórdenes en el aprendizaje de habilidades motoras así como desórdenes conductuales

### Lóbulo parietal

El **área somatosensorial primaria**, localizada en este lóbulo, recibe información sensorial que puede ser interpretada (dolor, temperatura, tacto, vibraciones, etc.). Similarmente al área motora primaria, la corteza sensorial primaria puede describirse mediante un homóculo sensorial (ver Figura 2.3, derecha). Una lesión en el lóbulo parietal puede producir diferentes y severos desórdenes, tales como trastorno de atención, de lenguaje, agnosia espacial (incapacidad de percibir los objetos normalmente), ataxia óptica (incapacidad de alcanzar con precisión un objeto), negligencia contralateral (incapacidad de diferencial izquierda y derecha), entre otros.

### Lóbulo temporal

Se compone principalmente por el **área auditiva primaria** (incluyendo el área de Wernicke), **área auditiva asociativa** y el **área temporal asociativa**, responsable de la memoria del lenguaje. Una lesión en el lóbulo temporal puede generar trastornos auditivos tales como sordera cortical y alucinaciones auditivas, entre otras. Asimismo trastornos en el lenguaje y en la capacidad de reconocer rostros puede ser consecuencia de lesiones en el lóbulo temporal.

### Lóbulo occipital

El **área visual primaria** y el **área visual asociativa** se encuentran en este lóbulo, en dónde la información captada por el órgano de la retina es procesada, permitiendo la interpretación de imágenes, el reconocimiento espacial y la discriminación del movimiento y colores. Una lesión a nivel del lóbulo occipital puede producir ceguera, distorsiones visual-perceptivas, en las que se persive una incapacidad para reconocer o comprender estímulos visuales (agnosia visual), daltonismo, alucinaciones visuales, entre otros.

El cerebelo es otro órgano fundamental del CNS. Este se encuentra localizado en la base del cerebro, y es el responsable de la coordinación de movimientos, el equilibrio y tono muscular. Un daño a nivel del cerebelo puede producir hipotónia, ataxia cerebelosa (descoordinación en el movimiento), trastornos motores, etc.



La médula espinal, localizada en el canal medular formado entre las vértebras, nace en la base del cráneo (foramen magnum) y se extiende hasta el sacro a lo largo de la columna vertebral. Los nervios espinales (también denominados raquídeos) son prolongaciones de la médula espinal que atraviesan los músculos vertebrales para distribuirse a las zonas del cuerpo, siendo estos denominados en función de su localización relativa a las vértebras. A la médula espinal llega información sensitiva (aférente) detectada por los receptores o nervios para ser retransmitida a las diversas áreas sensitivas del cerebro. Asimismo, desde la médula se emite información motora (eferente) desde el encéfalo a los nervios, haciendo que el organismo reaccione a los estímulos exteriores o interiores. Trastornos a nivel de la médula espinal producirán afecciones motoras y sensoriales acorde a la posición del daño.

### **2.1.2. Sistema nervioso periférico**

El sistema nervioso periférico (PNS, por sus siglas en inglés) está formado por todos los nervios periféricos que recorren el cuerpo, y se divide en el sistema nervioso somático y el sistema nervioso autónomo. El sistema somático es la parte del PNS responsable de llevar la información sensorial y motora hacia y desde el sistema nervioso central. Este sistema está formado por nervios de los receptores sensoriales del cuerpo, por ramas de nervios aferentes que llevan información de los receptores somáticos al CNS y por fibras eferentes que llevan los impulsos nerviosos del CNS al resto del cuerpo. Mientras que los nervios sensoriales están conformados por neuronas sensoriales, o también llamadas neuronas aferentes, que permiten recoger la información sensorial y enviarla al cerebro y a la médula espinal, los nervios motores están conformados por neuronas motoras, o también denominadas neuronas eferentes, que transportan la información del cerebro y la médula espinal a las fibras musculares de todo el cuerpo.

El sistema autónomo, el otro sistema que compone el PNS, es el responsable de regular las funciones corporales involuntarias, como el flujo sanguíneo, los latidos cardíacos, la digestión y la respiración.

## **2.2. Una nueva forma de comunicación**

Diferentes enfermedades pueden alterar o disminuir las vías de comunicación con las cuales el CNS controla su entorno. Enfermedades tales como la esclerosis lateral amiotrófica, accidente cerebro vascular, esclerosis múltiple, traumatismo medular o cerebral, parálisis cerebral, distrofia muscular, entre otras, dañan los canales neuronales que controlan

los músculos e incluso deterioran a los propios músculos (Wolpaw et al., 2002). La pérdida del control motor puede llegar a grados de severidad tales que la persona es incapaz de controlar voluntariamente su cuerpo, incluyendo el movimiento de ojos. Esto se denomina síndrome de enclaustramiento, enfermedad en que la persona es incapaz de comunicarse y controlar su propio cuerpo a pesar de tener sus capacidades cognitivas intactas.

En algunas situaciones no es posible revertir los efectos de estos desórdenes. Existen tres opciones para restablecer las funciones dañadas. Una primera alternativa es la de sustituir las vías dañadas por otras aún intactas (por ejemplo, un paciente puede comunicarse a través del movimiento ocular). La segunda opción busca restaurar la función dañada mediante la detección de actividad nerviosa o muscular por encima del nivel de la lesión. La tercera se basa en establecer una forma nueva y alternativa de interacción entre el cerebro de estas personas y su entorno (interno y externo) que prescinda de las vías neuromusculares normales. Las interfaces cerebro-computadora (BCIs, por sus siglas en inglés) son sistemas que miden la actividad cerebral del CNS y la convierten en señales artificiales de control que reemplazan, restablecen, aumentan, mejoran o suplen las salidas naturales del CNS (Wolpaw y McFarland, 1994). Este nuevo lenguaje de comunicación depende fuertemente del tipo de patrones provenientes de la actividad cerebral a utilizar para el control de un cierto dispositivo, así como de la adaptación mutua y continua tanto del sistema en sí como de su usuario final.

Como todo sistema de comunicación, un sistema BCI presenta una entrada (actividad cerebral), una salida (comandos de control), componentes que convierten (traducen) la entrada en salida, y un protocolo de operación que determina el comienzo, el final y los tiempos de la operación. Por ello, cualquier sistema BCI puede describirse mediante cuatro procesos fundamentales: i) adquisición de la señal, medición de la actividad cerebral mediante alguna tecnología; ii) procesamiento de señales, para traducir los patrones de activación cerebral en comandos de control; iii) dispositivo a controlar, que actúa según los comandos de control y refleja las intenciones del usuario; y iv) un protocolo de operación que guía el proceso (Schalk y Mellinger, 2010).

### 2.2.1. Midiendo la actividad cerebral

Existe una gran variedad de alternativas para monitorear la actividad cerebral, y en principio, proveer la señal de entrada a una BCI. Entre ellas podemos nombrar electroencefalografía (EEG), electrocorticografía, magnetoencefalografía, tomografía por emisión de positrones, imagen de resonancia magnética funcional y técnicas de infrarrojo cercano

funcional. Si bien cada técnica posee ciertas ventajas que la diferencian del resto, una forma no-invasiva, segura, rápida, y relativamente económica (comparada con los otros equipos de medición de la actividad cerebral), es mediante EEG de superficie. De hecho, es la tecnología predominante utilizada actualmente en los trabajos de BCI.

El EEG no-invasivo utiliza electrodos ubicados en el cuero cabelludo para medir la actividad eléctrica generada por las estructuras cerebrales, es decir mide la suma de los potenciales eléctricos post-sinápticos neuronales. Una característica relevante en el uso de EEG es la alta resolución temporal ( $< 1$  ms) que se obtiene a costa de una baja resolución espacial ( $\approx 2.5$  mm). De alguna manera, podríamos pensar que las imágenes cerebrales que podemos obtener mediante EEG responden al conocido principio de incertidumbre de Heisenberg: buena localización temporal conllevará a una mala localización espacial (Smith, 2004).

Para que la actividad cerebral sea observable mediante EEG de superficie debe producirse simultáneamente que:

- a) la mayor parte del potencial neuronal se produzca en una dirección perpendicular al cuero cabelludo,
- b) las dendritas neuronales estén dispuestas en paralelo, de modo que la suma de su campo potencial genere una señal detectable a una cierta distancia,
- c) la activación neuronal sea lo más sincrónica posible, y
- d) el signo de la actividad eléctrica producida por cada neurona sea el mismo.

Por otro lado la amplitud de los potenciales adquiridos por un electrodo cualesquiera depende de la intensidad y la distancia relativa a la fuente eléctrica, de la orientación espacial y de las propiedades eléctricas de las estructuras fisiológicas que separan al electrodo de la fuente.

Un método estándar y mundialmente aceptado para la distribución y posicionamiento de los electrodos a lo largo de toda la superficie del cuero cabello es el sistema internacional 10-20 (Jasper, 1958). Este sistema divide la cabeza en distancias proporcionales, utilizando como marcas el nasión, inion, mastoides y los puntos pre-auriculares. La Figura 2.4 muestra la configuración básica de este sistema con las distancias relativas entre electrodos y la denominación de los mismos según la posición que ocupan.

La actividad cerebral caracterizada por ciertos patrones puede ser asociada como la respuesta del CNS a ciertos estímulos o funciones, que pueden luego utilizarse como entradas de un sistema BCI. En los párrafos subsiguientes se describirán los paradigmas más utilizados en BCI, señalando sus características neurofisiológicas más relevantes.

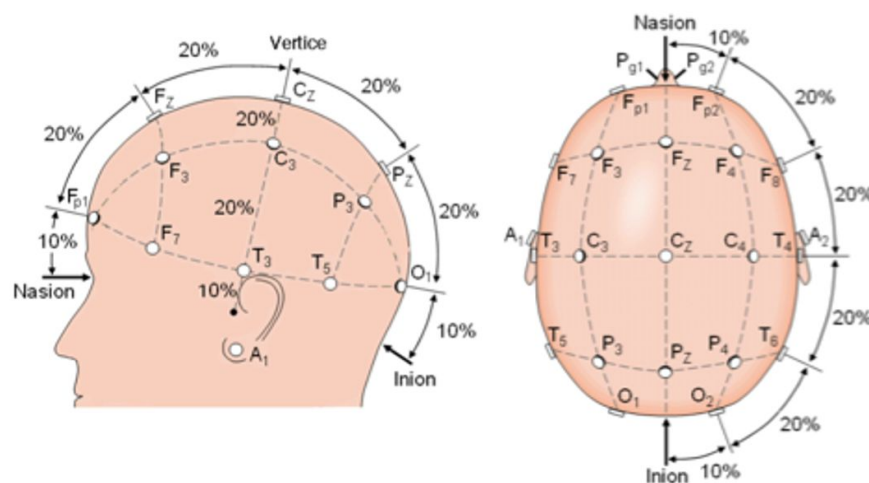


Figura 2.4: Vista lateral y superior del sistema internacional 10-20 de localización de electrodos junto con referencias de distancia y nombres de los electrodos más relevantes.

### 2.2.2. El lenguaje del cerebro

La señal de EEG puede dividirse en bandas frecuenciales que permiten la comprensión y análisis de ciertos comportamientos funcionales. La actividad cerebral, bajo ningún tipo de estímulo externo o interno, puede caracterizarse como la manifestación de oscilaciones impredecibles y aperiódicas con activaciones intermitentes que presentan picos espectrales en ciertas bandas observables. Por tal motivo la señal de EEG suele dividirse en cinco ritmos (bandas) principales: ritmos delta (0.1-3.5 Hz), ritmos theta (4-7.5 Hz), ritmos alfa (8-13 Hz), ritmos beta (14-30 Hz) y ritmos gamma (>30Hz).

Desde el punto de vista de la comunicación, un sistema BCI se basa en el uso de *símbolos* para comunicar la intención de su usuario, que luego la computadora deberá detectar y traducir. En particular estos símbolos o mensajes a detectar constituirán cierta información inmersa en la señal de EEG que varía según sea el paradigma de comunicación pre-seleccionado. En las BCI del tipo *reactivas*, el usuario envía los mensajes o comandos a la computadora en respuesta a un estímulo emitido por la interfaz. La interfaz luego “decide” qué tipo de comando se ha emitido por el usuario detectando eventualmente las modificaciones observables en la actividad cerebral (Clerc et al., 2016). En esta tesis abordamos las BCI del tipo reactivas. A continuación se describirán brevemente las características de dos de los paradigmas, también llamados protocolos de comunicación, más utilizados en BCI.

## BCI basadas en la onda P300

Cuando una persona recibe un estímulo externo (visual, auditivo o somatosensorial) inesperado y lo suficientemente significativo para que le llame la atención, se producirá como respuesta a nivel de CNS un potencial denominado potencial evocado relacionado a eventos (ERP, por sus siglas en inglés), el que puede ser considerado como el resultado de la reorganización de las fases en la señal de EEG en curso (Hillyard y Kutas, 1983). Los ERPs pueden describirse y diferenciarse por sus ondas componentes, es decir por la concatenación de una serie de respuestas al correspondiente estímulo. La denominada onda P300 es la componente más prominente y se caracteriza por poseer latencia mayor a 275 ms, polaridad positiva en los electrodos localizados en la línea media de la corteza, información frecuencial por debajo de los 12 Hz, y activaciones fuertemente observables en las áreas centrales y parietales del cerebro. La Figura 2.5 muestra la morfología de un ERP dadas sus ondas componentes. Si bien un ERP es una respuesta inconsciente, su amplitud está inversamente afectada por la probabilidad de ocurrencia (repetitividad) y directamente relacionada con la relevancia del estímulo (Fabiani et al., 1987).

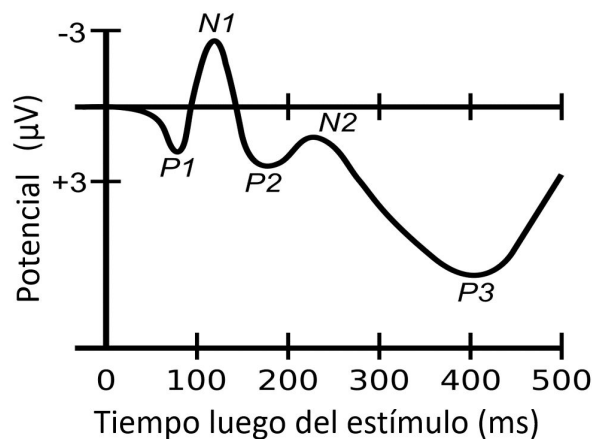


Figura 2.5: Señal ERP con sus ondas componente. Notar que, por convención, se adopta polaridad positiva hacia abajo.

La onda P300 representa un tipo de señal de entrada (marcador) muy utilizado en las BCI exógenas de tipo reactivas. Esto se debe a que: i) no se requiere entrenamiento previo del sujeto; ii) es una señal bien localizada en tiempo y de morfología conocida; y iii) el protocolo de comunicación (paradigma) permite una estimación rápida de dicha onda haciendo posible la comunicación. De hecho las BCI basadas en P300 pueden ser

comandadas por el 90 % de las personas (Wolpaw y Wolpaw, 2012).

Donchin y Farwell en 1988 (Farwell y Donchin, 1988) propusieron el uso de una BCI basada en la onda P300 como “prótesis mental” con el objetivo de deletrear palabras. Los autores se basaron en el paradigma “de bicho raro” (*oddball*, en inglés) aplicado a una matriz de  $6 \times 6$  con caracteres alfanuméricos. Durante el experimento, el sujeto es posicionado en frente de un monitor que presenta dicha matriz. A modo de estímulo, las filas y columnas de la matriz se iluminan o intensifican aleatoriamente. Dado que el objetivo final es deletrear palabras, el sujeto debe concentrarse en un caracter a la vez. Como labor mental debe contar la cantidad de veces que el caracter deseado es iluminado. Si la persona está debidamente concentrada, la iluminación del caracter deseado se transformará en un evento relevante y provocará un ERP como respuesta (Donchin et al., 2000). En cada bloque de intensificación, cada fila y cada columna son iluminadas sólo una vez. En consecuencia, hay 12 eventos posibles (6 filas y 6 columnas) de los cuales sólo 2 son relevantes (poseen P300). El posterior análisis de las señales permitirá identificar el caracter deseado como la intersección entre la fila y la columna que provocan dos P300 consecutivos.

### BCI basadas en ritmos sensoriomotores

En el caso de BCI endógenas, los cambios en la actividad cerebral deben ser inducidos por la misma persona, estando estos quizá determinados por una señal visual. La estimulación sensorial, la ejecución de movimientos así como la imaginación mental, producen cambios en la conectividad funcional en la corteza sensoriomotora que resultan en una disminución o en un incremento en la sincronización de las neuronas, y consecuentemente, una disminución de la amplitud (potencia) en ciertas bandas frecuenciales. A estas modulaciones se las conoce como eventos relacionados a la sincronización y desincronización de la actividad cerebral (ERS y ERD, por sus siglas en inglés, respectivamente) (Pfurtscheller y Aranibar, 1977). Existe una relación inversa entre frecuencia y amplitud de la señal, esto se evidencia en que, por ejemplo, oscilaciones de unos 10 Hz comprometen más neuronas sincronizadas que oscilaciones de 40 Hz (Singer, 1993).

La realización de un movimiento corporal genera cambios en la actividad cerebral observables mediante EEG de superficie en la región sensoriomotora. Dos segundos antes del comienzo de un movimiento en el hemisferio contralateral se presenta una predominante desincronización (patrones ERD) en las bandas mu (8-12 Hz) y beta (13-30) alcanzando su máximo bilateralmente con la ejecución del movimiento. Patrones de sincronización (ERS) alcanzan su máximo una vez finalizado el movimiento y suelen ser predominantes

en el lado ipsilateral (ver Figura 2.6). Asimismo también se ha observado un incremento rápido y de corta duración de la amplitud de los ritmos gamma (36-40 Hz) (Pfurtscheller y Da Silva, 1999). Esta ocurrencia simultánea de patrones ERD y ERS se conoce como ERD localizado/ERS circundante (Pfurtscheller y Neuper, 1997).

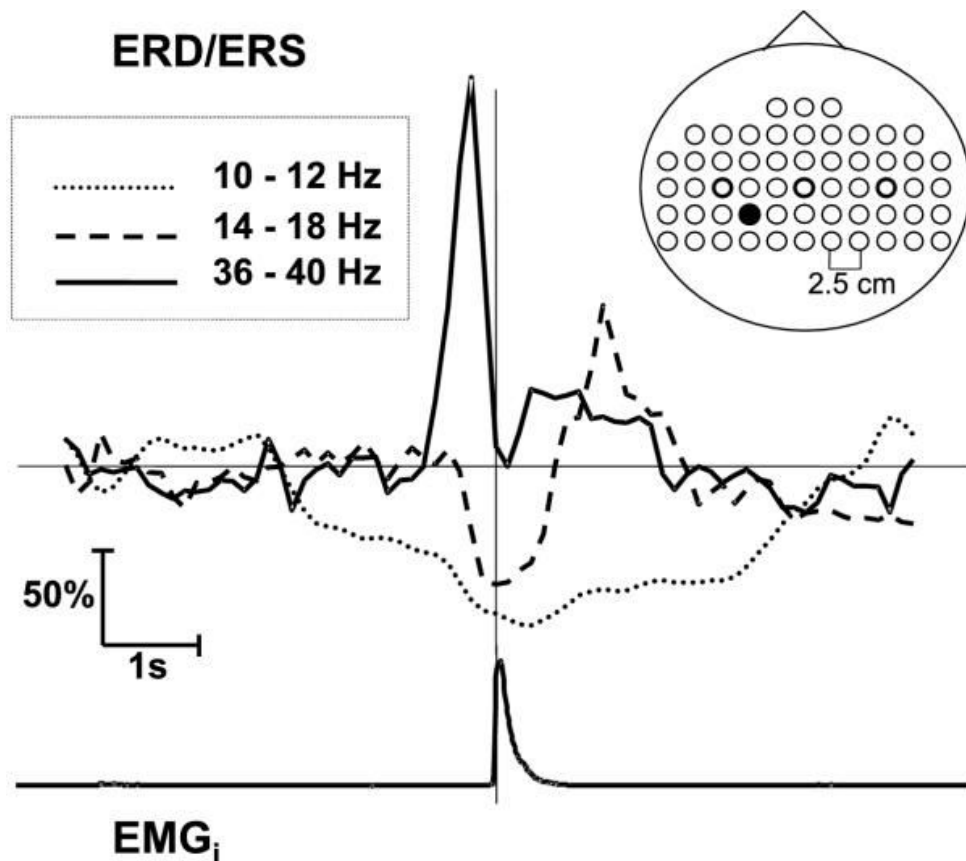


Figura 2.6: Patrones ERD/ERS observables antes, durante y después de la ejecución del movimiento, analizado en diferentes ritmos frecuenciales (10-12 Hz, 14-18 Hz y 36-40 Hz). La señal de EMG marca el comienzo del movimiento. Electrodo de análisis marcado en negro.

Existen tres tipos de estados fundamentales de la actividad cerebral (Pfurtscheller, 2001):

1. estado de reposo o neutro: sin ningún procesamiento específico de información sensorial, motora o cognitiva;
2. estado activado: con procesamiento aumentado de información en un sistema específico, mayor excitabilidad de las neuronas corticales y caracterizado por un ERD;

3. estado desactivado: con procesamiento de información reducido en un sistema específico, disminución de la excitabilidad de las neuronas corticales y caracterizado por un ERS en las bajas frecuencias.

Es bien sabido que mediante imaginería motora (MI, por sus siglas en inglés), es decir, mediante la simulación mental de la realización de ciertos movimientos, se producen modulaciones de menor amplitud en la actividad cerebral pero similares a las que se producen durante la ejecución del correspondiente movimiento. Asimismo, dado que la imaginación de diferentes partes del cuerpo produce patrones ERD/ERS con localización espacial bien determinados (debido al homóculo sensoriomotor descrito en la Figura 2.3), los ritmos sensoriomotores (SMRs, por sus siglas en inglés) son muy utilizados como señales de entrada en BCIs basadas en MI (Pfurtscheller y Neuper, 2001), siendo actualmente una técnica prometedora para terapias de rehabilitación de pacientes con pérdida severa del control voluntario de alguna de sus extremidades (Millán et al., 2010).

En BCI síncronas basadas en MI suele utilizarse un protocolo estándar de estimulación propuesto por el prestigioso grupo de investigación de Graz, Austria (Pfurtscheller y Neuper, 2001). Mediante este protocolo se le pide al sujeto que imagine mentalmente la realización de ciertos movimientos por unos pocos segundos (entre tres y cuatro) luego de la aparición de una señal visual. Imaginar un movimiento sin realmente ejecutarlo no es una tarea mental trivial, y la estrategia cognitiva llevada a cabo por el sujeto puede ser muy variable y difícil de caracterizar. Es por ello que en la práctica deben adaptarse y aprender tanto el sujeto como la BCI.

En la siguiente sección se realiza una breve descripción de las tareas involucradas en la decodificación de la actividad cerebral, eje central de esta tesis, así como un análisis del estado del arte y los desafíos actuales.

## 2.3. Detectando la actividad cerebral

El objetivo de los sistemas BCI basados en EEG, es identificar la actividad cerebral asociada a diferentes comandos de control. Para ello es necesario que el algoritmo de traducción sea capaz de aprender características en la señal de EEG que sean propias de la actividad cerebral asociada a una tarea específica, y suficientemente distintas a la actividad cerebral asociada a cualquier otro tipo de tarea. Desde el punto de vista de aprendizaje maquina, estamos ante un claro problema de reconocimiento de patrones.

En la Figura 1.1 se muestra esquemáticamente una BCI como diagrama en bloques. Nótese que luego del bloque de adquisición, la señal de EEG es generalmente tratada en



estadios separados y diferentes con el objeto de optimizar la comunicación. A continuación se realiza una breve descripción de cada bloque desde el punto de vista del procesamiento de señales.

### 2.3.1. Pre-procesamiento

El pre-procesamiento involucra la preparación de los registros de EEG. En esta etapa se suele filtrar y segmentar la señal, decisiones que impactarán directamente en el tratamiento futuro de la señal de EEG en los subsiguientes bloques. El filtrado y la segmentación temporal son utilizados para identificar y maximizar la información sobre ciertos rangos de tiempo y frecuencia asociados con la actividad cerebral a reconocer.

Las señales de EEG adquiridas mediante un amplificador sin ningún tipo de procesamiento se denominan señales crudas de EEG. Estas señales son series temporales, una por cada canal (electrodo), muestreadas a una cierta frecuencia pre-definida en el momento de la adquisición. El filtrado puede realizarse en el momento de la adquisición o posterior a este. En general, suele utilizarse un filtro pasa-banda aplicado a cada canal que elimine el ruido de línea u otros ruidos de alta frecuencia. Por otro lado, el proceso de segmentación extrae épocas o segmentos de EEG en una cierta ventana de tiempo, generalmente en la proximidad del comienzo del estímulo visual externo. Estos segmentos, también denominados observaciones o realizaciones desde el punto de vista estadístico, están asociados a un evento del protocolo, es decir a una clase del problema de reconocimiento de patrones (Clerc et al., 2016).

La segmentación resulta clave para el posterior procesamiento y clasificación de la señal, ya que luego de esta etapa tenemos un conjunto de observaciones (datos) cuya pertenencia a una de las clases del problema está definida por un vector de etiquetas.

En esta etapa también suelen removerse los artefactos de la señal, producidos principalmente por movimiento ocular, parpadeo y movimiento muscular de cejas y boca. Los artefactos contaminan la señal de EEG subyacente que uno espera elucidar para controlar un dispositivo externo. Si bien el uso de estas técnicas puede mejorar el tratamiento futuro de las señales de EEG, en este trabajo, bajo la premisa de construir métodos que sean robustos al ruido, no se eliminarán artefactos en la señal mediante técnicas de procesamiento de señales.

### 2.3.2. Extracción de características

El objetivo de la etapa de extracción de características es, precisamente, identificar de un conjunto de observaciones aquella información que es muy similar para una categoría, pero muy diferente para cualquier otra categoría (Duda et al., 2012). Para un sistema BCI se espera atribuir características o valores a los registros de EEG que sean representativos (discriminativos) para una clase y bien distintos con respecto a las otras clases. Estas características a aprender, claro está, dependerán del problema a tratar y las técnicas a utilizar deben tener en cuenta la naturaleza fisiológica de la señal utilizada como marcador en la BCI. Tres fuentes principales de información pueden utilizarse para extraer características de la señal de EEG (Clerc et al., 2016):

- Información espacial: describe de dónde (espacialmente) proviene la señal relevante.
- Información frecuencial: describe cómo la potencia de la señal de EEG varía en ciertas bandas de frecuencia específica.
- Información temporal: describe cómo la señal de EEG varía a lo largo del tiempo.

En el caso de los ERPs, por ser señales bien localizados en el tiempo, resulta natural utilizar la información temporal devenida por la variación de la señal a lo largo del tiempo. Asimismo la información espacial puede introducirse mediante el uso de aquellos electrodos ubicados principalmente en los lóbulos parietal y occipital, dado que es allí donde precisamente ocurre la mayor manifestación del P300, tal como se mencionó en 2.2.2. Las características de los ERPs suelen usualmente extraerse mediante tres simples pasos: i) filtrado pasa-banda entre 1-12 Hz., ii) submuestreo para reducir la dimensión del problema, y iii) concatenación de los canales para construir un único vector de características. Sin una selección apropiada del mínimo número de canales y muestras temporales, esto puede resultar en un problema de alta dimensionalidad cuyo tratamiento debe ser sobrellevado por el mismo clasificador.

Por otro lado, si la BCI utiliza SMRs, la característica más relevante a estudiar es la potencia de la señal de EEG en ciertas bandas frecuenciales (información frecuencial) y en ciertas regiones cerebrales (información espacial). Así, un diseño básico de BCI-basadas en MI podría utilizar información espacial mediante la extracción de características de electrodos ubicados en las áreas motoras, esencialmente C3 y C4. Asimismo, podría utilizar información espectral focalizándose en los ritmos mu y beta al calcular la banda de potencia en dichas frecuencias. Esto resulta en un vector de características de tan sólo cuatro valores. Como desventaja, la información de unos pocos electrodos es utilizada en bandas de frecuencia prefijadas.

Ambos diseños presentan soluciones sencillas aunque altamente mejorables. Diferentes métodos de extracción de características se han propuesto, los cuales discutiremos en más detalles en los capítulos subsecuentes.

### 2.3.3. Selección de características

La selección de características es una etapa opcional que previene la acumulación de información irrelevante para describir la actividad de EEG relacionada a un comando específico. Demasiadas características podrían provocar una generalización pobre por parte del clasificador, aumentar la complejidad computacional y requerir mayor cantidad de patrones de entrenamiento para alcanzar un cierto desempeño. Para esta etapa suelen utilizarse enfoques de Teoría de la Información que aporten información cuantitativa sobre la relevancia de ciertas características sobre otras, identificando aquellas con mayor poder de discriminación y descartando aquellas otras con poco o nulo aporte al problema de clasificación. Medidas de discrepancia entre distribuciones, información mutua, entre otras, son las más utilizadas.

En el marco de esta tesis nos focalizaremos en la utilización de información a-priori de la discrepancia entre clases apoyándonos en Teoría de la Información.

### 2.3.4. Clasificación

La tarea del clasificador es transformar el conjunto de características en un conjunto de salidas de comandos de control. Desde el punto de vista de aprendizaje estadístico, un clasificador es una función  $f(x) : \mathbb{R}^d \rightarrow \mathcal{Y}$ , la cual dada una observación  $\mathbf{x} \in \mathbb{R}^d$  predice el valor de la etiqueta  $y \in \mathcal{Y}$ . Notar que el valor de  $y$  puede ser categórico o nominal. Las coordenadas o valores del vector  $\mathbf{x}$  son las características extraídas de una observación  $x$ , tal como se explicó en la Sección [2.3.2](#). En la práctica, la función  $f(\cdot)$  es estimada dado un conjunto de  $n$  ejemplos (patrones) de *entrenamiento*  $\{\mathbf{X}, \mathbf{y}\} = \{(\mathbf{x}_i, y_i), \mathbf{x}_i \in \mathbb{R}^d, y_i \in \mathcal{Y}, i = 1, \dots, n\}$ , y se espera que dado un conjunto de futuras observaciones aún no vistas (denominado patrones de *testeo*) sea capaz de predecir correctamente la etiqueta de clase.

Cuán exitosa sea la etapa de clasificación está determinado principalmente por tres criterios: precisión, eficiencia computacional y complejidad. Un algoritmo de traducción eficiente debe ser capaz de obtener buenos resultados de clasificación, poseer buena velocidad de comunicación y mínima complejidad de la solución ([Li et al., 2011a](#); [Wolpaw et al., 2002](#)). Asimismo, para BCIs del tipo individuales, el modelo aprendido (algoritmo

de traducción) que comprende la extracción, selección y clasificación de características, debe ser capaz de seleccionar aquella información que mejor represente la intención del sujeto (sujeto-dependiente), mediante técnicas que sean capaces de adaptarse a los posibles cambios de la señal de EEG, para así proveer salidas de control efectivas para ser utilizadas en aplicaciones en tiempo real.

En esta tesis se persigue el objetivo de aliviar la tarea de clasificador buscando soluciones que contemplen e integren la selección de características basadas en información discriminativa a-priori.

## 2.4. Estado del arte y desafíos actuales

Las señales de EEG utilizadas en sistemas BCIs no invasivos fluctúan fuertemente a lo largo del tiempo (ver Sección [2.2.1](#) para comprender qué mide cada canal de EEG), no sólo entre diferentes sesiones y sujetos, sino también dentro de una misma sesión y persona. Estas últimas, denominadas variaciones intra-sujeto, suelen deberse tanto a factores técnicos, como la posición e impedancia de los electrodos, como a factores humanos, como la fatiga, disminución de la motivación o cambios físicos, etc. Las variaciones entre sujetos son aún más difíciles de tratar, ya que tienen su origen en diferencias anatómicas, funcionales y psicológicas. Estas variabilidades en la señal de EEG definen, claramente, a la señal de EEG como una señal de tipo no-estacionaria ([Shenoy et al., 2006](#)). El nivel de ruido es también un gran obstáculo a superar. Los denominados artefactos de la señal pueden ser de tipo fisiológico, como la actividad muscular, movimiento de ojos y parpadeo, así como no-fisiológicos tal como el ruido de línea (50 Hz). Sumado a esto, la señal de interés suele ser de unos pocos  $\mu V$  mientras que el EEG en curso puede llegar a valores de  $mV$ , produciendo una muy baja relación señal-ruido (SNR, por sus siglas en inglés).

Asimismo los algoritmos de traducción deben enfrentarse a la conocida “maldición de la dimensionalidad”. Por trabajar con señales temporales provenientes de un gran número de canales, la dimensionalidad de los vectores es más grande que la cantidad total de observaciones (ver Subsección [2.3.2](#)). Dado que se recomienda entrenar con una cantidad de observaciones que superen, en al menos, cinco a diez veces la dimensión de los patrones por cada clase, y como, desafortunadamente, en los sistemas BCI la cantidad de datos para entrenar suele ser limitada, estamos ante una clara “maldición” ([Lotte et al., 2007](#)).

Por los motivos expuestos, los sistemas BCI basados en técnicas de clasificación tradicionales, requieren de una larga sesión de calibración para poder recolectar una cantidad

de patrones lo suficientemente grande para entrenar el algoritmo de traducción de manera confiable (Zhang et al., 2013). A pesar de que estas sesiones resultan largas y demandantes para el usuario, permiten, a su vez, que este sea capaz de familiarizarse con el sistema y adquirir la experiencia necesaria para aprender a usar una BCI. Por lo tanto, al diseñar una BCI debe balancearse la complejidad tecnológica para interpretar la actividad cerebral con la cantidad de entrenamiento necesario para obtener una comunicación exitosa (Allison et al., 2012). Por este motivo y para aumentar la practicabilidad de los sistemas BCI, construir un algoritmo de decodificación en escenarios de entrenamiento pequeño que sea lo suficientemente robusto, en el sentido de su capacidad de arrojar resultados similares ante cambios en los datos, y lo suficientemente confiable, en el sentido de producir los mismo niveles de clasificación de manera repetitiva, es aún un gran desafío.

Las técnicas actuales de decodificación de la actividad cerebral suelen utilizar y mejorar métodos estándar de aprendizaje maquinal (Lotte et al., 2007). La interdisciplinariedad característica de las BCI ha llevado a que actualmente los métodos de extracción, selección y clasificación de características utilicen técnicas que han mostrado eficiencia en otras aplicaciones de procesamiento de señales y reconocimiento de patrones. Los métodos de representación ralas de señales han cobrado gran interés en los últimos años debido a sus grandes avances tanto teóricos como prácticos en diversas aplicaciones (Huang y Aviyente, 2007; Li et al., 2014). Mediante soluciones basadas en representación rala, la dimensionalidad de un problema puede ser reducida drásticamente a uno pocos coeficientes relevantes distintos de cero, se puede aumentar la velocidad de computo y almacenamiento, así como obtener soluciones con mayor grado de generalización. Por otro lado, la incorporación de métodos de regularización para mejorar las soluciones de problemas “mal condicionados” ha sido también foco de atención en los últimos años (Lotte et al., 2018; Tomioka y Müller, 2010). Clasificadores basados en técnicas de regularización suelen proveer soluciones más robustas a valores atípicos y con mejores propiedades de generalización, aún en escenarios poco favorables como lo son los de entrenamiento pequeño (pocos patrones de entrenamiento de alta dimensionalidad).

La introducción de información a-priori en el modelo discriminativo para guiar la selección de las características más relevantes para la discriminación de clases puede involucrar la combinación del conocimiento neurofisiológico del problema y de información acumulada en los datos mismos. Como se mencionó en 2.3.3 medidas bien conocidas de Teoría de la Información pueden utilizarse para cuantificar dicha información. Además, es importante mencionar que el pre-procesamiento, la extracción de características y la clasificación no deben atacarse como problemas aislados, puesto que de este modo se ig-

noraría la interacción entre dichos procesos, resultando en desempeños sub-óptimos (Mak et al., 2011).

Finalmente, como se indica en (Wolpaw y Wolpaw, 2012) un sistema BCI ideal debe ser seguro, asequible, estable, confiable, que no requiera de la asistencia permanente de un técnico o investigador, que restaure la comunicación a una velocidad “normal”, y que sea estéticamente aceptable. Varios de estos requerimientos implican ciertas exigencias en el algoritmo de traducción (tales como las mencionadas más arriba), mientras que otras están directamente ligadas al sistema de medición de la actividad cerebral. Uno de los mayores desafíos actuales en la comunidad de BCI es el de poder transferir las BCI más allá del ámbito de la investigación, de modo de poder acercar las BCI a sus usuarios finales. Una de las principales limitaciones en este sentido se debe a los elevados costos para acceder al equipamiento necesario para adquirir las señales de EEG (con una alta relación señal-ruido), que sea transportable y fácil de usar.

Tal como se mencionó en el capítulo introductorio, en esta tesis se han abordado varios de estos desafíos aquí planteados, con el objetivo principal de desarrollar algoritmos de traducción que sean lo suficientemente confiables (precisión por encima del 70% (Li et al., 2011b)), robustos y veloces en la decodificación de la intención del usuario; para así obtener sistemas BCI que puedan luego trasladarse fácilmente a escenarios de la vida real



## Capítulo 3

# Datos reales: registros de EEG en BCI

“Existían obstáculos concretos, pero la vida consiste precisamente en aceptar dichos obstáculos.”

---

Franz Kafka, La metamorfosis

A lo largo de este trabajo diferentes base de datos se han utilizado para validar la eficiencia de los métodos propuestos, así como para analizar las ventajas y desventajas de los mismos en relación a otros métodos del estado del arte. Se han dividido y referenciado las bases de datos según sea el paradigma utilizado en la construcción de las mismas.

### 3.1. Base de datos de EEG bajo el paradigma P300

Las bases de datos utilizadas a lo largo de este trabajo para detección de ERPs, se han seleccionado por ser bases de datos muy utilizadas en la literatura. Por otro lado, dado que una de las base de datos cuenta con una gran cantidad de patrones adquiridos mediante un número muy grande de electrodos para sólo *dos* sujetos de experimentación, mientras que la otra posee registros de EEG de *varios* sujetos de experimentación pero con una menor cantidad de patrones y canales, es de interés analizar el comportamiento de los algoritmos en estos dos bien diferenciados escenarios muestrales. Para la construcción de ambas bases de datos, se ha utilizado el paradigma de “bicho raro” junto con la matriz de Donchin y Farwell descrita en la Sección [2.2.2](#). A continuación se describen brevemente ambas bases de datos.



### 3.1.1. P300-1: Base de datos ERP-EEG del LINI-UAM, México

El conjunto de datos P300-1 está conformado por una base de datos de acceso público desarrollada por el “Laboratorio de Investigación en Neuroimagenología de la Universidad Autónoma Metropolitana”, Ciudad de México, México. Esta base de datos consiste en registros de EEG de 25 sujetos sanos adquiridos mediante 10 electrodos (Fz, C3, Cz, C4, P3, Pz, P4, PO7, PO8, y Oz) localizados según el sistema internacional 10-20. Se utilizó el amplificador gUSBamp (g.tec, Austria) con una frecuencia de muestreo de 256 Hz. Una matriz de  $6 \times 6$  con caracteres alfanuméricos fue utilizada durante el experimento para deletrear palabras. Como estímulo visual, cada fila y cada columna de la matriz era iluminada por un período de 62.5 ms, con un intervalo inter-estímulo de 125 ms. Por cada carácter a deletrear se repitió el bloque de estimulación 15 veces (Ledesma-Ramirez et al., 2010).

Cada sujeto participó en cuatro (4) sesiones, de las cuales las dos primeras eran de deletreo dirigido, es decir, se le indicó al sujeto qué palabras debía deletrear. Dado que estas sesiones poseen el valor verdadero del vector de etiquetas, nuestro conjunto de datos P300-1 está construido con registros de EEG provenientes de las dos primeras sesiones. Cada sujeto debió deletrear en total 21 caracteres. Dado que de las 12 intensificaciones (6 filas y 6 columnas), sólo dos provocan un evento relevante (ERP), y dado que cada bloque de intensificación fue repetido 15 veces, para cada sujeto de la base de datos P300-1 se tiene un total de  $21 \times 12 \times 15 = 3780$  realizaciones de EEG, de las cuales sólo 630 contienen ERP (clase objetivo).

Como pre-procesamiento, las señales fueron filtradas entre 0,1 y 12 Hz mediante un filtro de orden 4 Butterworth pasa-banda. Segmentos de EEG de 1000 ms fueron extraídos inmediatamente después de la presentación del estímulo. La dimensión final de los patrones es  $10 \times 256 = 2560$ .

### 3.1.2. P300-2: Base de datos II de la Competencia III de BCI

Esta base de datos, que denotaremos a lo largo del trabajo como P300-2, corresponde a la base de datos II de la Competencia III de BCI<sup>1</sup>. Esta base de datos contiene registros de EEG de dos sujetos sanos (A y B) muestreados a 240 Hz utilizando 64 canales localizados según el sistema 10-20. Los datos se encuentran divididos en *entrenamiento* y *testeo*, en los cuales los sujetos debían deletrear 85 y 100 caracteres, respectivamente. Cada fila y cada columna de la matriz de caracteres alfanuméricos fue intensificada durante 100 ms

<sup>1</sup><http://www.bbci.de/competition/iii/>

con un tiempo inter-estimulo de 75 ms. Cada bloque de intensificación fue repetido 15 veces (Blankertz et al., 2006).

A lo largo de esta tesis sólo los datos de entrenamiento fueron utilizados para conformar el conjunto de datos P300-2. Esta elección se debe a que estos registros contienen el valor verdadero de las etiquetas y además la cantidad de realizaciones es suficientemente grande ( $85 \times 12 \times 15 = 15300$ , de las cuales sólo 2550 contienen ERP).

Idéntico pre-procesamiento que para el conjunto de datos P300-1 fue realizado para esta base de datos. Asimismo, para disminuir la dimensionalidad de los patrones, y conociendo cuales son las áreas cognitivas más relevantes en la generación del P300, se seleccionaron 16 de los 64 canales (F3, Fz, F4, T7, C3, Cz, C4, T8, P7, P3, Pz, P4, P8, PO7, PO8, y Oz) tal como en los trabajos (Zhang et al., 2014, 2013). En consecuencia, la dimensión de cada realización es de  $16 \times 240 = 3840$ .

## 3.2. Base de datos de EEG bajo el paradigma de MI

En este trabajo hemos utilizado tres bases de datos de EEG bajo el paradigma de imaginación motora, dos de las cuales son bases de datos de acceso público de las Competencias de BCI<sup>2</sup>, mientras que la tercera es una base de datos propia, adquirida por la tesista durante su estancia doctoral en el “Rehabilitation Engineering Laboratory”, ETH, Zurich, Suiza.

### 3.2.1. MI-1: Base de datos propia

Esta base de datos corresponde a los registros de 11 sujetos sanos sin experiencia previa en BCI (3 mujeres, 4 de dominancia izquierda, edad promedio  $25,45 \pm 2,50$  años) adquiridos en la primer sesión del experimento detallado en el Capítulo 7. Dichos experimentos fueron realizados por la tesista en el “Rehabilitation Engineering Laboratory”, ETH, Zurich, Suiza. Las señales fueron adquiridas utilizando 64 electrodos superficiales localizados en concordancia con el sistema internacional 10-20. Se utilizó como referencia y tierra el electrodo CPz y AFz, respectivamente. Las señales fueron muestreadas a 512 Hz y filtradas entre 0,5 Hz y 40 Hz con un filtro de 3<sup>er</sup> orden Butterworth pasa-banda por el software de adquisición. Luego estas fueron submuestreadas a 128 Hz. Inspirado en el uso de BCIs para rehabilitación motora funcional, el protocolo involucraba dos tipos de tareas: imaginación del gesto de abrir y cerrar la mano dominante desde una perspecti-

---

<sup>2</sup><http://www.bbc.de/competition/>

va propia y sensorial (cinestética) y relajación/no-movimiento. Para validar el protocolo, electromiografía (EMG) de superficie de dos músculos de la mano y de dos músculos del antebrazo fueron adquiridos en simultáneo con el EEG (Noraxon TeleMyo DTS, Noraxon, Estados Unidos).

La sesión estaba compuesta de cuatro rondas o corridas. Cada corrida de experimentación constaba de 20 repeticiones de imaginación motora cinestética (de 4 s de duración) intercaladas aleatoriamente con 20 épocas de no imaginación/relajación por el mismo período de tiempo. En total, para cada sujeto se registraron 160 épocas de EEG (80 de cada clase). La máxima contracción voluntaria (MVC, por sus siglas en inglés), es decir la máxima fuerza que una persona puede realizar en un ejercicio isométrico específico, fue adquirida al comienzo de la sesión. El análisis de la señal de EMG mostró que uno de los sujetos de esta base de datos contenía un porcentaje muy grande de épocas de EEG contaminadas por actividad muscular. Por este motivo, dicho sujeto fue desestimado para el análisis de estos datos, quedando en total, registros de EEG de sólo 10 sujetos sanos.

Más detalles sobre este protocolo de estimulación, el procesamiento de la señal de EMG y el registro de los datos, pueden encontrarse en el Capítulo [7](#).

### 3.2.2. MI-2: Base de datos IVa de la Competencia III de BCI

Este conjunto de datos está conformado por la base de datos IVa de la Competencia III de BCI realizada en 2006. Esta base de datos contiene registros de EEG de 5 sujetos sanos (aa, al, av, aw, ay) adquirida mediante 118 electrodos ubicados según el sistema internacional 10-20 extendido. En el momento de la adquisición las señales fueron filtradas entre 0,05 y 200 Hz. Las señales registradas a 1000 Hz, fueron luego submuestreadas a 100 Hz. Cada sujeto realizó 280 repeticiones de imaginación motora de su mano derecha y de su pie derecho (140 para cada condición). Las tareas de MI fueron indicadas mediante señales visuales y debían realizarse por un período de 3,5 s ([Blankertz et al., 2006](#)).

### 3.2.3. MI-3: Base de datos IIb de la Competencia IV de BCI

La base de datos IIb de la Competencia IV de BCI corresponde a registros de EEG de 9 sujetos sanos de dominancia derecha de un estudio publicado en ([Leeb et al., 2007](#)). Tres canales bipolares fueron registrados (C3, Cz y C4) con una frecuencia de muestreo de 250 Hz. Cada sujeto se sometió a 5 sesiones. En cada sesión los sujetos debían realizar MI de su mano izquierda o derecha por un periodo de 4,5 s.

Nuestro conjunto de datos, denominado como MI-3, está conformado por los datos

---

correspondientes a la tercera sesión (B0103T, B0203T,..., B0903T), tal como en (Thomas et al., 2009; Zhang et al., 2015b). Este conjunto de datos contiene 160 realizaciones de EEG (80 para cada clase de MI) por sujeto.

Es importante mencionar que para las bases de datos bajo el paradigma MI las clases se encuentran balanceadas, es decir la cantidad de patrones provenientes de cada clase es la misma; mientras que para las bases de datos que involucran el paradigma basado en P300, el desbalance entre clases es muy grande, teniendo señales de EEG sin P300 (clase no-objetivo) y señales de EEG con P300 (clase objetivo) en una relación 10/2.



## Capítulo 4

# Discriminación rala y regularización con penalizantes mixtos

“Pensar es olvidar diferencias, es generalizar, abstraer.”

---

Jorge Luis Borges, Funes el memorioso.

En este capítulo presentaremos el método principal desarrollado en esta tesis. Dicho método, en el cuál hacemos uso de regularización con penalizantes mixtos, nace luego de un análisis del estado del arte tanto en el área de BCI como en el de estadística aplicada. Presentaremos primero una breve revisión de los antecesores del método, para luego introducir los formalismos e implementaciones del mismo.

### 4.1. Análisis discriminante lineal: su concepción y evolución

El análisis discriminante lineal (LDA, por sus siglas en inglés) es una herramienta muy utilizada en diferentes áreas para reducción de dimensionalidad y clasificación. Su popular uso se debe principalmente a que otorga soluciones robustas, es sencillo de implementar y tiene bajo costo computacional (Clemmensen et al., 2012).

Sean  $\mathbf{W}_1, \dots, \mathbf{W}_K$  vectores aleatorios  $p$ -dimensionales cuyas distribuciones caracterizan unívocamente a cada una de las  $K$  clases de un cierto problema de clasificación. Sean además  $\mathbf{X}^{n \times p}$ ,  $\mathbf{y} \in \mathbb{R}^n$ , tales que cada una de las filas de  $\mathbf{X}$  es una realización de uno y sólo uno de los vectores aleatorios antes mencionados y  $y_i = k$ , si y sólo si la  $i$ -ésima fila de  $\mathbf{X}$  proviene de una realización de  $\mathbf{W}_k$  (es decir  $k$  es la clase a la que “pertenece” la  $i$ -ésima

fila de  $\mathbf{X}$ ). Nos referiremos a la matriz  $\mathbf{X}$  como “matriz de datos”, conformada por  $n$  “observaciones”  $\mathbf{x}_i$  y al vector  $\mathbf{y} \in \{1, 2, \dots, K\}^n$  como el vector de etiquetas categóricas.

En problemas de clasificación el objetivo general es estimar el vector de etiquetas  $\mathbf{y}$  a partir de  $\mathbf{X}$ . El método LDA genera una transformación lineal  $\boldsymbol{\beta} \in \mathbb{R}^{p \times q}$  que mapea las observaciones  $\mathbf{x}_i$  de  $\mathbb{R}^p$  en  $\mathbb{R}^q$ , donde  $q < K$ . Para hallar  $\boldsymbol{\beta}$  se deben calcular la matriz de covarianza intra-clases, la matriz de covarianza entre-clases y la matriz de covarianza total, definidas respectivamente por:

$$\boldsymbol{\Sigma}_w = \frac{1}{n} \sum_{k=1}^K \sum_{i \in I_k} (\mathbf{x}_i - \boldsymbol{\mu}_k)(\mathbf{x}_i - \boldsymbol{\mu}_k)^T, \quad (4.1)$$

$$\boldsymbol{\Sigma}_b = \frac{1}{n} \sum_{k=1}^K n_k (\boldsymbol{\mu}_k - \boldsymbol{\mu})(\boldsymbol{\mu}_k - \boldsymbol{\mu})^T, \quad (4.2)$$

$$\boldsymbol{\Sigma}_t = \frac{1}{n} \sum_{i=1}^n (\mathbf{x}_i - \boldsymbol{\mu})(\mathbf{x}_i - \boldsymbol{\mu})^T, \quad (4.3)$$

donde  $I_k$  y  $n_k$  denotan el conjunto de índices y la cantidad de patrones pertenecientes a la  $k$ -ésima clase,  $\boldsymbol{\mu}_k \doteq \frac{1}{n_k} \sum_{i \in I_k} \mathbf{x}_i$  es la media de la clase  $k$  y  $\boldsymbol{\mu} \doteq \frac{1}{n} \sum_{k=1}^K n_k \boldsymbol{\mu}_k$  es la media común a todas las clases. Notar que  $\boldsymbol{\Sigma}_t = \boldsymbol{\Sigma}_w + \boldsymbol{\Sigma}_b$ .

La transformación  $\boldsymbol{\beta}$  busca simultáneamente maximizar  $\boldsymbol{\Sigma}_b$  y minimizar  $\boldsymbol{\Sigma}_w$ , que equivale a simultáneamente maximizar  $\boldsymbol{\Sigma}_b$  y minimizar  $\boldsymbol{\Sigma}_t$ , es decir:

$$\boldsymbol{\beta} = \arg \max_{\boldsymbol{\beta} \in \mathbb{R}^{p \times q}} (\boldsymbol{\beta}^T \boldsymbol{\Sigma}_b \boldsymbol{\beta}) (\boldsymbol{\beta}^T \boldsymbol{\Sigma}_t \boldsymbol{\beta})^{-1}. \quad (4.4)$$

Dado que  $\boldsymbol{\Sigma}_b$  tiene como máximo rango  $K - 1$ , hay a lo sumo  $K - 1$  soluciones no triviales para el problema (4.4), y por lo tanto habrán a lo sumo  $K - 1$  vectores discriminativos. Estos vectores son precisamente las direcciones sobre las cuales los datos presentan máxima covarianza entre-clase en relación a su covarianza intra-clase. Usualmente se calculan  $q = K - 1$  vectores solución.

Cuando la cantidad de observaciones es menor a la dimensión de las mismas ( $n < p$ ), la matriz  $\boldsymbol{\Sigma}_t$  resulta generalmente mal condicionada, lo que produce una mala estimación de  $\boldsymbol{\beta}$  y, consecuentemente, una deficiente clasificación.

En el caso particular  $K = 2$  la solución de (4.4) tiene la siguiente forma explícita:

$$\boldsymbol{\beta} = \boldsymbol{\Sigma}_t^{-1} (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2). \quad (4.5)$$

En este caso estamos en presencia del conocido Análisis Discriminante Lineal de Fisher (FLDA, por sus siglas en inglés) el que puede formularse como un problema de regresión lineal (Duda et al., 2012; Mika, 2002), (ver demostración en Anexo A.1).

Varios autores (e.g. (Hastie et al., 1994), (Ye, 2007), (Hastie et al., 2009), como para nombrar algunos) han extendido esta formulación para problemas multiclase, en los que se ha demostrado una estrecha relación entre LDA y regresión multivariada. En particular, en 1998 Hastie et. al (Hastie et al., 1994) introducen una formulación no-lineal (más flexible) del análisis discriminante al que denominan “puntaje óptimo” (*optimal scoring*, en inglés), el cual describiremos brevemente a continuación.

Sea  $\mathbf{X}$  como antes y  $\mathbf{Y}$  una matriz de  $n \times K$  de variables binarias, donde  $Y_{ij}$  es un indicador de si la observación  $i$ -ésima pertenece a la clase  $j$ -ésima. Definimos  $\Theta = [\theta_1, \dots, \theta_q] \in \mathbb{R}^{K \times q}$ , donde los  $\theta_j$  se obtienen recursivamente para  $j = 1, 2, \dots, q$ , como la solución del problema de *optimal scoring*:

$$\begin{aligned} (\beta_j, \theta_j) &= \arg \min_{\beta \in \mathbb{R}^p, \theta \in \mathbb{R}^K} \|\mathbf{Y}\theta - \mathbf{X}\beta\|_2^2, \\ \text{s.a. } &\frac{1}{n} \theta^T \mathbf{Y}^T \mathbf{Y} \theta = 1, \quad \theta^T \mathbf{Y}^T \mathbf{Y} \theta_l = 0 \quad \forall l = 1, 2, \dots, j-1. \end{aligned} \quad (4.6)$$

Para  $j = 1$  la condición de ortogonalidad en (4.6), la cual se impone a los efectos de evitar soluciones triviales, desaparece. Detalles de la implementación computacional para resolver (4.6) pueden encontrarse en (Hastie et al., 1994).

Nos referiremos a cada  $\theta_j$  como un vector de “puntajes”. Observar que  $\theta_j$  es el vector en  $\mathbb{R}^K$  para el cual la transformación  $\mathbf{Y} \rightarrow \mathbf{Y}\theta_j$  resulta óptima para el problema de mínimos cuadrados implícitamente definido en (4.6). Es precisamente este mapeo,  $\mathbf{Y} \rightarrow \mathbf{Y}\theta_j$ , el que introduce mayor flexibilidad al análisis discriminante, puesto que transforma una matriz binaria en una matriz con componentes reales arbitrarios.

En 2011, Clemmensen et. al (Clemmensen et al., 2012) proponen una versión regularizada del problema de *optimal scoring* llamado análisis discriminante ralo (SDA, por sus siglas en inglés) mediante la adición de dos términos de penalización en (4.6), lo que permite incluir variables correlacionadas e inducir rareza en la solución. Más precisamente, el problema SDA consiste en resolver recursivamente para  $j = 1, 2, \dots, q$ , el siguiente



problema:

$$\begin{aligned}
 (\boldsymbol{\beta}_j, \boldsymbol{\theta}_j) &= \arg \min_{\boldsymbol{\beta} \in \mathbb{R}^p, \boldsymbol{\theta} \in \mathbb{R}^K} \{ \|\mathbf{Y}\boldsymbol{\theta} - \mathbf{X}\boldsymbol{\beta}\|_2^2 + \lambda_1 \|\boldsymbol{\beta}\|_1 + \lambda_2 \|\boldsymbol{\beta}\|_2^2 \}, \\
 \text{s.a. } & \frac{1}{n} \boldsymbol{\theta}^T \mathbf{Y}^T \mathbf{Y} \boldsymbol{\theta} = 1, \quad \boldsymbol{\theta}^T \mathbf{Y}^T \mathbf{Y} \boldsymbol{\theta}_l = 0 \quad \forall l = 1, 2, \dots, j-1,
 \end{aligned} \tag{4.7}$$

donde  $\lambda_1$  y  $\lambda_2$  son parámetros positivos prefijados denominados parámetros de regularización que controlan el grado de penalización. Más adelante analizaremos formas apropiadas para la selección de estos parámetros.

El problema (4.7) se resuelve de manera iterativa y alternada hasta satisfacer algún criterio de convergencia. Primero fijamos  $\boldsymbol{\theta}_j$  y optimizamos con respecto a  $\boldsymbol{\beta}_j$  y luego optimizamos  $\boldsymbol{\theta}_j$  con  $\boldsymbol{\beta}_j$  fijo, es decir, se iteran hasta alcanzar algún criterio de convergencia los siguientes dos pasos:

1. Para  $\boldsymbol{\theta}_j$  fijo, resolver:

$$\boldsymbol{\beta}_j = \arg \min_{\boldsymbol{\beta} \in \mathbb{R}^p} \{ \|\mathbf{Y}\boldsymbol{\theta}_j - \mathbf{X}\boldsymbol{\beta}\|_2^2 + \lambda_1 \|\boldsymbol{\beta}\|_1 + \lambda_2 \|\boldsymbol{\beta}\|_2^2 \}. \tag{4.8}$$

2. Para  $\boldsymbol{\beta}_j$  fijo, resolver:

$$\begin{aligned}
 \boldsymbol{\theta}_j &= \arg \min_{\boldsymbol{\theta} \in \mathbb{R}^K} \|\mathbf{Y}\boldsymbol{\theta} - \mathbf{X}\boldsymbol{\beta}_j\|_2^2, \\
 \text{s.a. } & \frac{1}{n} \boldsymbol{\theta}^T \mathbf{Y}^T \mathbf{Y} \boldsymbol{\theta} = 1, \quad \boldsymbol{\theta}^T \mathbf{Y}^T \mathbf{Y} \boldsymbol{\theta}_l = 0 \quad \forall l = 1, 2, \dots, j-1.
 \end{aligned}$$

Para más detalles sobre la implementación ver (Clemmensen et al., 2012; Sjöstrand et al., 2012).

La solución del problema (4.7), tal como en LDA, genera  $q$  direcciones discriminativas,  $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \dots, \boldsymbol{\beta}_q$ , sobre las cuales las clases de la matriz de datos proyectados  $(\mathbf{X}\boldsymbol{\beta}_1 \ \mathbf{X}\boldsymbol{\beta}_2 \ \dots \ \mathbf{X}\boldsymbol{\beta}_q) \in \mathbb{R}^{n \times q}$  pueden ser bien distinguidas mediante un sencillo clasificador lineal.

El problema planteado en (4.8), conocido como “red elástica” (*Elastic-Net*, ENET, por sus siglas en inglés) (Zou y Hastie, 2005), es similar al operador de mínima reducción y selección absoluta (LASSO, por sus siglas en inglés) (Tibshirani, 1996), ya que no sólo realiza selección rala de variables, sino que además permite la selección de variables correlacionadas (debido al término de penalización en la norma  $\ell_2$ ).

## 4.2. Análisis discriminante lineal ralo generalizado

Existe una gran variedad de posibles métodos para resolver problemas de clasificación. Cada uno de estos métodos requiere del cumplimiento de ciertas hipótesis, cuyo cumplimiento debe verificarse a partir de los datos disponibles. Es razonable suponer que cuanto más información a-priori pueda ser incorporada (de manera apropiada) en un método de clasificación, más eficiente será la discriminación entre clases. Por lo tanto, es importante tratar de extraer la mayor cantidad de información contenida en los datos sobre el problema a atacar.

Sean  $\mathbf{X}$ ,  $\mathbf{Y}$ ,  $\beta_j$ ,  $\theta_j$ ,  $\lambda_1$  y  $\lambda_2$  como antes, y sean  $\mathbf{D}_1$  y  $\mathbf{D}_2$  matrices diagonales definidas positiva de dimensión  $p \times p$  que cuantifican apropiadamente la diferencia entre clases. El método, denominado análisis discriminante ralo generalizado (GSDA, por sus siglas en inglés), consiste en resolver para  $K = 2$  el siguiente problema de regularización con restricciones:

$$\begin{aligned} (\hat{\beta}, \hat{\theta}) &= \arg \min_{\beta \in \mathbb{R}^p, \theta \in \mathbb{R}^K} \{ \|\mathbf{Y}\theta - \mathbf{X}\beta\|_2^2 + \lambda_1 \|\mathbf{D}_1\beta\|_1 + \lambda_2 \|\mathbf{D}_2\beta\|_2^2 \}, \\ \text{s.a.} \quad &\frac{1}{n} \theta^T \mathbf{Y}^T \mathbf{Y} \theta = 1, \end{aligned} \quad (4.9)$$

Al igual que antes, el problema (4.9) se resuelve de manera iterativa y alternada, es decir:

1. Dado  $\theta$ , solución de (4.11), resolver:

$$\hat{\beta} = \arg \min_{\beta \in \mathbb{R}^p} \{ \|\mathbf{Y}\theta - \mathbf{X}\beta\|_2^2 + \lambda_1 \|\mathbf{D}_1\beta\|_1 + \lambda_2 \|\mathbf{D}_2\beta\|_2^2 \}. \quad (4.10)$$

2. Dado  $\beta$  solución de (4.10), resolver:

$$\hat{\theta} = \arg \min_{\theta \in \mathbb{R}^2} \|\mathbf{Y}\theta - \mathbf{X}\beta\|_2^2 \quad \text{s.a.} \quad \frac{1}{n} \theta^T \mathbf{Y}^T \mathbf{Y} \theta = 1. \quad (4.11)$$

El vector  $\hat{\beta}$ , obtenido como solución de (4.9), hereda las propiedades de rareza e inclusión de variables correlacionadas de SDA y, además, contiene información discriminativa en cada  $p$ -ésima muestra con el objeto de mejorar la separabilidad entre clases. Como antes, la clasificación se realiza sobre la matriz  $\mathbf{X}\hat{\beta}$  de dimensión  $n \times 1$ .

A continuación presentaremos formas para construir apropiadamente las matrices  $\mathbf{D}_1$  y  $\mathbf{D}_2$  de modo que las mismas incluyan información discriminativa entre clases.

## Matrices de anisotropía

Una manera de incluir información discriminativa en las matrices de anisotropía  $\mathbf{D}_1$  y  $\mathbf{D}_2$  es mediante alguna cuantificación de la diferencia entre las distribuciones de probabilidad entre clases. Si bien existe una gran variedad de “métricas” para comparar distribuciones de probabilidad (Basseville, 1989), aquí utilizaremos la conocida divergencia de Kullback-Leibler. La decisión de utilizar esta métrica no sólo se debe a sus propiedades matemáticas, sino también a su popular uso para la selección de características en problemas de clasificación en diferentes aplicaciones (Gersch et al., 1979; Gupta et al., 2009; Moreno et al., 2003). En el Capítulo 5 se analizará el impacto de utilizar diferentes métricas.

Sea  $\mathbf{N}$  una variable aleatoria discreta definida en el espacio discreto  $\mathcal{N}$  y sean  $f_1(n)$  y  $f_2(n)$ ,  $n \in \mathcal{N}$ , dos funciones de probabilidad. Luego, la divergencia de Kullback-Leibler (KLD, por sus siglas en inglés) de  $f_1$  relativa a  $f_2$  está definida como:

$$d_{\text{KL}}(f_1||f_2) \doteq \sum_{n \in \mathcal{N}} f_1(n) \log \left( \frac{f_1(n)}{f_2(n)} \right), \quad (4.12)$$

con la convención  $0 \cdot \log 0 \doteq 0$ . Si bien  $d_{\text{KL}}(f_1||f_2)$  mide la discrepancia entre  $f_1$  y  $f_2$ , no es una distancia en el sentido matemático estricto, ya que no es simétrica y no satisface la desigualdad triangular. En este contexto es preferible contar con una medida simétrica, tal como la divergencia-J, la que se define como:

$$d_J(f_1, f_2) \doteq \frac{d_{\text{KL}}(f_1||f_2) + d_{\text{KL}}(f_2||f_1)}{2}. \quad (4.13)$$

Denotemos ahora con  $f_k^i(\cdot)$  a la función de probabilidad de la clase  $k$  en la característica  $i$ , con  $k = 1, 2$  y  $i = 1, 2, \dots, p$ . Definimos entonces la divergencia-J en la característica  $i$  como

$$d_J(i) \doteq d_J(f_1^i, f_2^i). \quad (4.14)$$

Esta función así definida cuantifica la discrepancia entre dos clases en la característica  $i$ . Un valor de  $d_J(i)$  cercano a cero significa que en la  $i$ -ésima característica hay poca o nula información de discriminación entre clases, mientras que un valor grande de  $d_J(i)$  implica que la característica  $i$  contiene información relevante sobre la discrepancia entre clases. Por lo tanto, dado que sería sumamente deseado poder resaltar (y seleccionar) aquellas características que sean relevantes en la clasificación y evitar la selección de aquellas características no discriminativas, las matrices  $\mathbf{D}_1$  y  $\mathbf{D}_2$  se definen de la siguiente

manera:

$$\begin{aligned}\mathbf{D}_1 &\doteq \text{diag}(1 - \alpha_i + \alpha_i c_i), \\ \mathbf{D}_2 &\doteq \text{diag}(c_i),\end{aligned}$$

donde

$$c_i \doteq \frac{\left(\prod_{j=1}^p d_J(j)\right)^{1/p}}{d_J(i)}, \quad \alpha_i \doteq \frac{\max\{c_j\}_{j=1}^p - c_i}{\max\{c_j\}_{j=1}^p - \min\{c_j\}_{j=1}^p}, \quad i = 1, \dots, p.$$

Notar que con  $\mathbf{D}_1$  y  $\mathbf{D}_2$  así definidas,  $c_i$  es grande cuando  $d_J(i)$  es chico, y vice-versa. El parámetro  $\alpha_i$  ( $0 \leq \alpha_i \leq 1$ ,  $\forall i = 1, \dots, p$ ) pesa la información de KLD proporcionalmente a su relevancia. De esta manera  $\alpha_i = 1$  si  $d_J(i) = \max\{d_J(j)\}_{j=1}^p$  y  $\alpha_i = 0$  si  $d_J(i) = \min\{d_J(j)\}_{j=1}^p$ . Notar que si existe  $i_0$ ,  $1 \leq i_0 \leq p$ , tal que  $d_J(i_0) = 0$ , entonces la definición formal de las matrices  $\mathbf{D}_1$  y  $\mathbf{D}_2$  carece de sentido. Esta situación, de todos modos, puede ser fácilmente sobrellevada reemplazando  $d_J(i_0)$  por  $d_J(i_0) + \epsilon$ , con  $\epsilon \gtrsim 0$  pequeño.

### Implementación computacional

Para implementar GSDA, se utilizó como base el algoritmo SDA original propuesto en (Clemmensen et al., 2012). Por lo tanto, GSDA es básicamente resuelto en dos pasos. El primer paso requiere resolver (4.10), que es una generalización del problema ENET (Mouret et al., 2013). El segundo paso consiste en actualizar el vector de puntajes  $\boldsymbol{\theta}$  resolviendo (4.11).

Para resolver el primer paso, podemos re-formular (4.10) como un problema de mínimos cuadrados de tipo LASSO. Para ello definimos,

$$\tilde{\mathbf{X}} \doteq \begin{pmatrix} \mathbf{X} \\ \sqrt{\lambda_2} \mathbf{D}_2 \end{pmatrix}_{(n+p) \times p}, \quad \tilde{\mathbf{Y}} \doteq \begin{pmatrix} \mathbf{Y}\boldsymbol{\theta} \\ \mathbf{0}_{p \times 1} \end{pmatrix}_{(n+p) \times 1},$$

y el problema de ENET generalizado (4.10) puede entonces re-escribirse como:

$$\hat{\boldsymbol{\beta}} = \arg \min_{\boldsymbol{\beta} \in \mathbb{R}^p} \{\|\tilde{\mathbf{Y}} - \tilde{\mathbf{X}}\boldsymbol{\beta}\|_2^2 + \lambda_1 \|\mathbf{D}_1 \boldsymbol{\beta}\|_1\}, \quad (4.15)$$

el cual es conocido como LASSO generalizado (Tibshirani y Taylor, 2011). Si  $\mathbf{D}_1$  es invertible, la solución de (4.15) puede explícitamente calcularse como  $\hat{\boldsymbol{\beta}} = \mathbf{D}_1^{-1} \hat{\boldsymbol{\alpha}}$ , donde  $\hat{\boldsymbol{\alpha}}$

es la solución de:

$$\hat{\boldsymbol{\alpha}} = \arg \min_{\boldsymbol{\alpha} \in \mathbb{R}^p} \{ \|\tilde{\mathbf{Y}} - \tilde{\mathbf{X}}\mathbf{D}_1^{-1}\boldsymbol{\alpha}\|_2^2 + \lambda_1 \|\boldsymbol{\alpha}\|_1 \}. \quad (4.16)$$

De esta manera con la ecuación hallada en (4.16), podemos utilizar el conocido algoritmo LARS-EN (Zou y Hastie, 2005) en la resolución de (4.10). Este algoritmo es un método iterativo que selecciona en cada paso aquella variable que esté más correlacionada con el residuo, hasta alcanzar algún criterio de parada. Tiene la ventaja de ser computacionalmente tan costoso como un simple LSR. Para la implementación de GSDA, hemos utilizado y modificado el algoritmo LARS-EN descrito en (Sjöstrand et al., 2012). Más detalles sobre la implementación de GSDA se presentan en la siguiente subsección.

### Selección de los parámetros de regularización

Es claro que la elección de los parámetros de regularización es clave para obtener la mejor solución posible del problema (4.9). El criterio de la Curva-L (Hansen, 1992) es una técnica popular para la elección de parámetros de regularización dado que no requiere de ningún conocimiento a-priori de la norma del error. La Curva-L es una gráfica de tipo log-log de la norma del residuo en función de la norma de la solución. De esta manera, la Curva-L muestra el compromiso entre la minimización de estas dos cantidades. Se sabe que el vértice o esquina de la Curva-L corresponde al punto donde la regularización y el error se encuentran balanceados. Este punto se define como aquel de máxima curvatura y como se muestra en (Belge et al., 1998) coincide con el mínimo del residuo. El método originalmente fue diseñado para encontrar el parámetro de regularización para problemas de regularización de tipo Tikhonov-Phillips (Hansen, 1992).

En (Belge et al., 1998) se propone extender el método de la Curva-L para la selección simultánea de parámetros de regularización en problemas de regularización mixta. Este método es denominado Hipersuperficie-L. En el contexto de GSDA formalizado en (4.9) o más precisamente en el contexto del problema ENET generalizado (4.10), la hipersuperficie-L se define como  $S(\boldsymbol{\lambda}) \doteq \{(x_1(\boldsymbol{\lambda}), x_2(\boldsymbol{\lambda}), z(\boldsymbol{\lambda})) : \boldsymbol{\lambda} \in \mathbb{R}_+^2\}$ , donde  $x_1(\boldsymbol{\lambda}) \doteq \log \|\mathbf{D}_1\boldsymbol{\beta}(\boldsymbol{\lambda})\|_1$ ,  $x_2(\boldsymbol{\lambda}) \doteq \log \|\mathbf{D}_2\boldsymbol{\beta}(\boldsymbol{\lambda})\|_2^2$  y  $z(\boldsymbol{\lambda}) \doteq \log \|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}(\boldsymbol{\lambda})\|_2^2$ . Luego, el vector óptimo de parámetros de regularización está definido por  $\hat{\boldsymbol{\lambda}} \doteq \arg \min_{\boldsymbol{\lambda} \in \mathbb{R}_+^2} z(\boldsymbol{\lambda})$ .

A pesar de que ENET generalizado está definido mediante el uso de  $(\lambda_1, \lambda_2)$ , este no es la única opción para la selección de los parámetros de ajuste (Zou y Hastie, 2005). Dada la existente relación entre LASSO y ENET demostrada en (4.16), se puede utilizar la fracción de la norma- $\ell_1$  ( $t = \frac{\|\boldsymbol{\beta}\|_1}{\max \|\boldsymbol{\beta}\|_1} \in [0, 1]$ ) como opción para  $\lambda_1$  o la cantidad de elementos distintos de cero de la solución ( $|\boldsymbol{\beta}|$ ). Asimismo, dado que LARS-EN es un

algoritmo iterativo, la cantidad de pasos ( $\kappa$ ) también puede ser utilizada como parámetro de ajuste en lugar de  $\lambda_1$ . En este trabajo elegimos como par de parámetros de ajuste a  $(\lambda_2, \kappa)$ , es decir,  $\boldsymbol{\lambda} \doteq (\lambda_2, \kappa)$ . Mediante la definición  $z(\boldsymbol{\lambda}) \doteq \|\mathbf{Y}\boldsymbol{\theta} - \mathbf{X}\boldsymbol{\beta}(\boldsymbol{\lambda})\|_2^2$ , y dado que los puntos donde la curvatura de la superficie es máxima coinciden con los puntos donde la superficie del error es mínima (Belge et al., 1998), el vector de parámetros óptimos  $\boldsymbol{\lambda}^*$  fue seleccionado como aquel que minimiza la norma residual. El algoritmo de GSDA, junto con la selección automática de parámetros, se describe en Algoritmo 1. El paquete de funciones del algoritmo GSDA es de código abierto y de libre acceso, disponible en <https://github.com/vpeterson/gsd>.

### 4.3. Experimentos y resultados

Para analizar el efecto de la incorporación de la información a-priori en el problema de minimización con penalización mixta es necesario comparar el desempeño obtenido por el método propuesto GSDA versus el método original SDA. Para dicha comparación se utilizaron las bases de datos P300-1 y P300-2 descritas en el Capítulo 3, subsecciones 3.1.1 y 3.1.2, respectivamente. Estas bases de datos corresponden a 3780 épocas de EEG (630 con P300, 3150 sin P300) de dimensión 2560, y a 15300 épocas (2550 con P300, 12750 sin P300) de dimensión 3840, respectivamente. Notar que  $n > p$ .

#### 4.3.1. KLD y matrices de anisotropía

Un análisis de la divergencia-J como función de la información espacial (canales) y la información temporal (muestras temporales) nos permite detectar cuáles son aquellas características más discriminativas en cada canal e instante de tiempo. La Figura 4.1 presenta tres gráficas en el plano tiempo-canal a modo ilustrativo para los sujetos 1 y 14 de P300-1, y para ambos sujetos (A y B) de la P300-2. Estas gráficas muestran la información sobre los valores de KLD y de las matrices  $\mathbf{D}_1$  y  $\mathbf{D}_2$  construidas en función de KLD.

Las gráficas de KLD en la Figura 4.1, indican que la mayor cantidad de la información discriminativa está localizada en una ventana de tiempo entre 250 y 500 ms, lo cuál concuerda con la latencia de la onda P300 (ver Capítulo 1, Sección 2.2.2). Por otro lado, se puede observar, y con mayor grado en los sujetos de P300-2, que hay canales que aparentemente no contribuyen a la separación entre clases. En el caso del Sujeto-14 de P300-1, la gráfica de KLD aparenta no tener una región discriminativa clara y precisa

**Algorithm 1** GSDA con selección automática de parámetros

**Inputs:**  $\mathbf{X}$ ,  $\mathbf{Y}$ ,  $\mathbf{D}_1$ ,  $\mathbf{D}_2$ ,  $\Lambda_2 = \{\lambda_2^{(1)}, \dots, \lambda_2^{(d)}\}$ .

- 1: Definir  $\boldsymbol{\pi} \doteq \frac{1}{n} \mathbf{Y}^T \mathbf{Y}$ ,  $K = 2$
- 2: Inicialización:  $\boldsymbol{\theta} = \text{eye}(K, 1)$ .
- 3: **while** hasta alcanzar algún criterio de parada **do**
- 4:     **for**  $i = 1, \dots, d$  **do**
- 5:         Re-definición de las variables:

$$\tilde{\mathbf{X}} = \begin{pmatrix} \mathbf{X} \\ \sqrt{\lambda_2^{(i)}} \mathbf{D}_2 \end{pmatrix}_{(n+p) \times p}, \quad \tilde{\mathbf{Y}} = \begin{pmatrix} \mathbf{Y}\boldsymbol{\theta} \\ \mathbf{0}_{p \times 1} \end{pmatrix}_{(n+p) \times 1}$$

- 6:     Resolver el problema ENET generalizado:

$$(\mathbf{A}, \kappa) = \text{LARSEN}(\tilde{\mathbf{X}}\mathbf{D}_1^{-1}, \tilde{\mathbf{Y}}), \quad \mathbf{B} = \mathbf{D}_1^{-1}\mathbf{A}$$

- 7:     Hallar el residuo:

$$\mathbf{R}(\lambda_2, 1 : \kappa) = \|\mathbf{Y}\boldsymbol{\theta} - \mathbf{X}\mathbf{B}_j\|_2^2$$

- 8:     Guardar la solución:

$$\mathbf{B}_{\text{all}}(\lambda_2, 1 : \kappa, :) = \mathbf{B}$$

- 9:     **end for**
- 10:     Seleccionar la dirección óptima:

$$(\hat{\lambda}_2, \hat{\kappa}) = \arg \min_{\lambda_2, \kappa} \mathbf{R}(\lambda_2, \kappa)$$

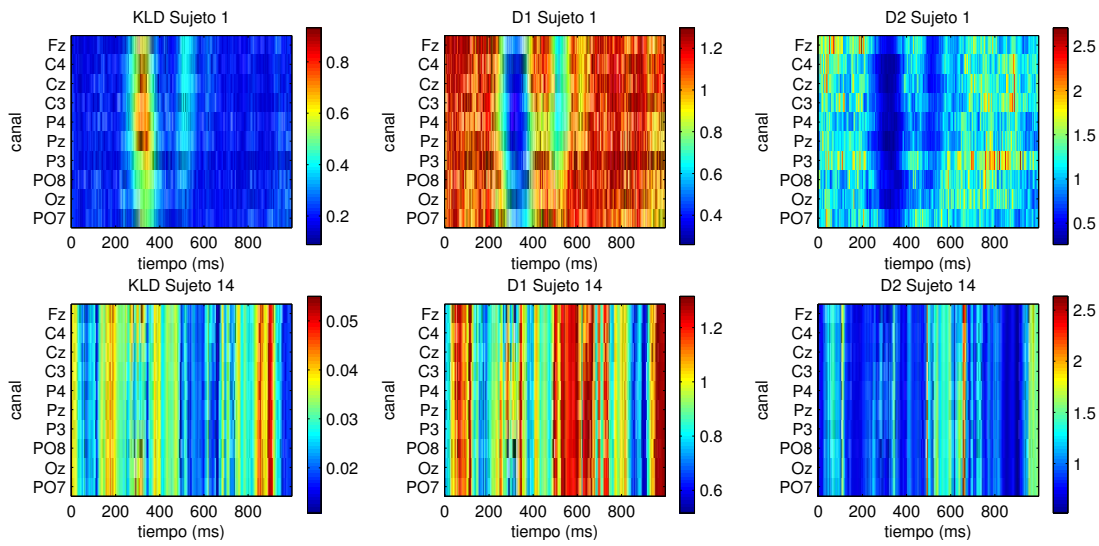
$$\boldsymbol{\beta} = \mathbf{B}_{\text{all}}(\hat{\lambda}_2, \hat{\kappa}, :)$$

- 11:     Actualizar  $\boldsymbol{\theta}$ :

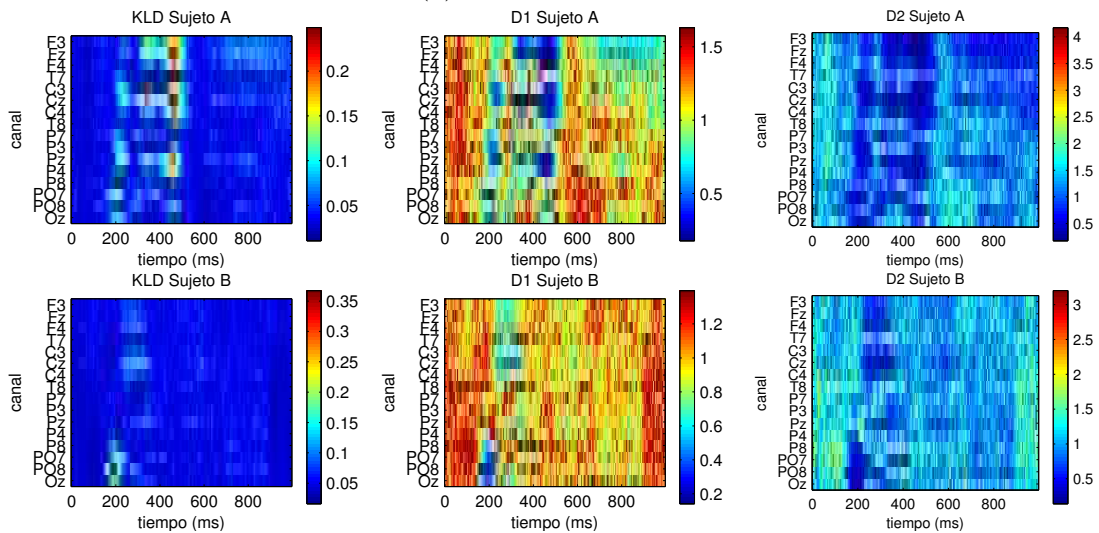
$$\tilde{\boldsymbol{\theta}} = (\mathbf{I} - \boldsymbol{\theta}\boldsymbol{\theta}^T \boldsymbol{\pi}) \boldsymbol{\pi}^{-1} \mathbf{Y}^T \mathbf{X} \boldsymbol{\beta}, \quad \boldsymbol{\theta} = \frac{\tilde{\boldsymbol{\theta}}}{\sqrt{\tilde{\boldsymbol{\theta}}^T \boldsymbol{\pi} \tilde{\boldsymbol{\theta}}}}$$

- 12: **end while**

**Outputs:**  $\boldsymbol{\theta}$ ,  $\boldsymbol{\beta}$



(a) Sujeto 1 y 14 de la P300-1



(b) Sujeto A y B de la P300-2

Figura 4.1: Divergencia-J (KLD) y matrices de anisotropía ( $D_1$  y  $D_2$ ) para diferentes sujetos de las bases de datos P300-1 y P300-2.

asociada al P300. Nótese también cómo en las gráficas de  $D_1$  y  $D_2$  no se penalizan aquellos puntos en lo que KLD es grande.



### 4.3.2. Desempeño de GSDA vs. SDA

#### Resultados de clasificación

Para la implementación numérica de SDA ( $\mathbf{D}_1 = \mathbf{D}_2 = \mathbf{I}$ ) y de nuestro método GSDA, se utilizó el algoritmo LARS-EN modificado tal como se describió en Algoritmo 1. Se hizo variar al parámetro  $\lambda_2$  entre  $10^{-6}$  y  $10^{-1}$  en escala logarítmica. Para poder comparar SDA y GSDA bajo la misma condición de parada, se fijó la cota superior de la norma  $\ell_1$  para P300-1 y P300-2 igual a 10 % y 20 % de la dimensión original de los patrones ( $p$ ), respectivamente.

Los experimentos fueron realizados utilizando validación cruzada de 3 particiones fijas, es decir los métodos fueron siempre evaluados considerando la misma partición de datos en cada sujeto. En cada partición la información a-priori medida por la divergencia-J (4.14) fue estimada en los datos de entrenamiento (ver Figura 4.1). El vector solución  $\hat{\beta}$  hallado tanto para SDA o GSDA era luego utilizado para proyectar la matriz de datos de entrenamiento, en el cual un simple clasificador lineal (LDA) puede luego construirse. El desempeño de clasificación alcanzado por cada método fue evaluado mediante el área bajo la curva ROC (AUC, por sus siglas en inglés) (Bradley, 1997), dada su ventaja en el análisis de problemas de clases fuertemente desbalanceados (Fawcett, 2006). El análisis de significancia estadística de los resultados de clasificación arrojados por cada método fue realizado mediante ANOVA un factor con bloques aleatorizados, con nivel de significancia  $\alpha = 0,05$ .

Los resultados de clasificación de SDA y GSDA para la base de datos P300-1 y P300-2, se muestran en la Figura 4.2 y la Tabla 4.1, respectivamente.

Tabla 4.1: Promedio y desvío estándar de los valores del área bajo la curva ROC (AUC) alcanzados por SDA y GSDA en los datos de testeo para la base de datos P300-2 evaluada mediante validación cruzada de 3 particiones.

	SDA	GSDA
<b>Sujeto A</b>	0,7547 ( $\pm$ 0,0087)	<b>0,7551</b> ( $\pm$ 0,0098)
<b>Sujeto B</b>	0,8418 ( $\pm$ 0,0041)	<b>0,8490</b> ( $\pm$ 0,0024)
<b>Promedio</b>	0,7982 ( $\pm$ 0,0033)	<b>0,8020</b> ( $\pm$ 0,0053)

Tal como se puede ver en la Figura 4.2 y en la Tabla 4.1, GSDA supera a SDA para ambos sujetos de la base de datos P300-2 y para todos los sujetos salvo *uno* de los 25 sujetos de la base de datos P300-1. Estos resultados muestran que la apropiada inclusión

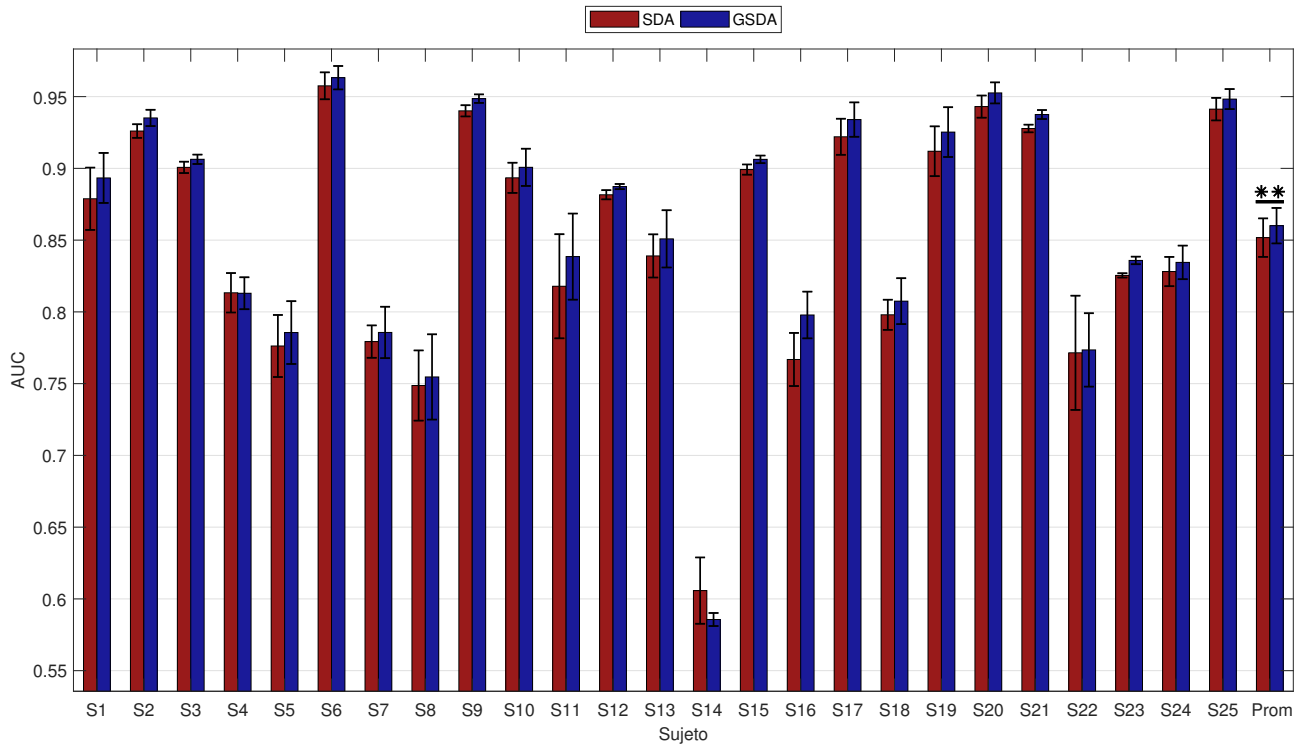


Figura 4.2: Valores del área bajo la curva ROC (AUC) alcanzados en los datos de testeo al utilizar SDA y GSDA para la base de datos P300-1 evaluada mediante validación cruzada de 3 particiones. Para cada sujeto, las barras y las líneas negras representan el promedio y el desvío estándar, respectivamente, de los valores de AUC en las tres particiones. La última columna corresponde al promedio del desempeño de todos los sujetos. Aquí “\*\*” indica que  $GSDA > SDA$  con  $p\text{-valor} < 0,05$ .

de la información discriminativa a-priori en el modelo puede influenciar positivamente los resultados de clasificación. En el caso del Sujeto-14 de P300-1, la inclusión de la información KLD impacta negativamente en el desempeño de clasificación. Esto puede deberse, tal como se discutió en [4.3.2](#), a que aparentemente la información discriminativa para este sujeto no se presenta en una región localizada asociada al P300, haciendo que

la información introducida en el modelo pueda quizá deberse a ruido, lo cual resulta en última instancia perjudicial para la clasificación.

### Análisis de la rareza de la solución

Una solución con mayor grado de rareza es una solución con mayor grado de robustez. Si este grado de rareza está acompañado en una mejora en los resultados de clasificación, entonces la herramienta de selección de características utilizada en el modelo resulta óptima. Para ello, se estudió el número final de variables distintas de cero del vector solución respecto a la dimensión original de los patrones para ambos métodos ( $\frac{|\hat{\beta}|}{p} \%$ ). Para la base de datos P300-1, se encontró que, en promedio, este porcentaje era de alrededor del 14 % y 6 %, para SDA y GSDA, respectivamente. Para la base de datos P300-2, estos valores fueron del 24 % y 5 %, respectivamente. Estos resultados muestran que GSDA no sólo mejora la clasificación de SDA sino que además requiere de menos del 6 % de la dimensión original de los patrones para alcanzar dicho desempeño, resultando en soluciones más robustas y confiables.

### Costo computacional

El costo computacional asociado a la cantidad de iteraciones requeridas por cada método hasta hallar el vector solución  $\hat{\beta}$  fue también analizado. La Figura 4.3 muestra para cada base de datos, el promedio de la cantidad de iteraciones finales requeridas por cada método hasta encontrar el vector solución. Como se puede observar, nuestro método GSDA converge en un menor número de iteraciones en comparación con las requeridas por SDA. Esto, consecuentemente, se traduce en una disminución de los tiempos computacionales. En nuestros experimentos (MATLAB®, PC Intel® Core™ i7-6700K CPU @ 4.00GHz  $\times$  8 con 32GB de memoria), encontramos que cuando se utiliza GSDA el tiempo de corrida del algoritmo fue disminuido en promedio en más de un 2,9 % y en más de un 9,4 % comparado con SDA para la base de datos P300-1 y P300-2, respectivamente. Estos resultados muestran que la penalización mediante la información KLD “guía” mejor al algoritmo a su convergencia al evitar la selección de aquellas variables que no son relevantes en términos de discriminabilidad.

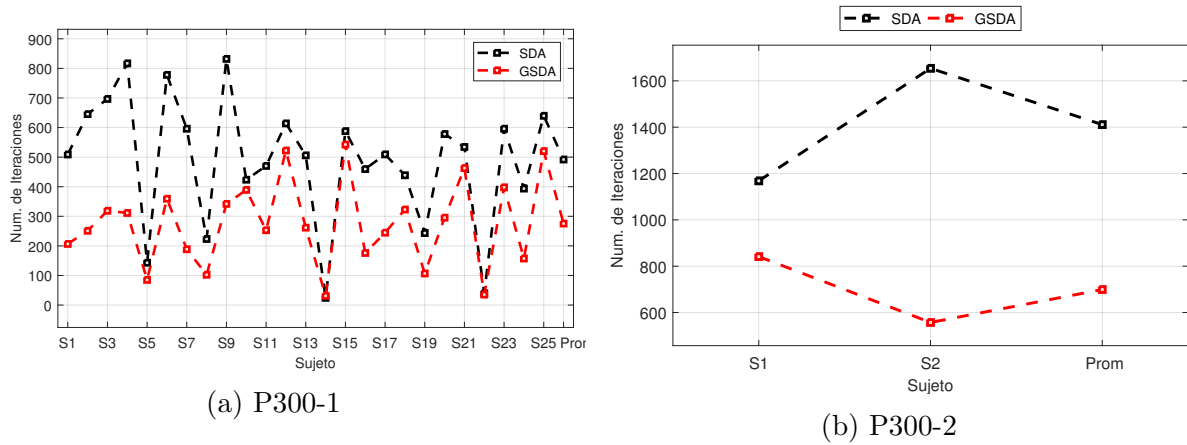


Figura 4.3: Cantidad de iteraciones requeridas por SDA y GSDA para encontrar el vector solución  $\hat{\beta}$  para la base de datos P300-1 (a) y P300-2 (b).

## 4.4. Discusiones

A la luz de los resultados expuestos en las subsecciones anteriores podemos concluir que la incorporación de la información a-priori sobre la discrepancia entre clases resulta beneficiosa para la clasificación, disminuyendo los tiempos de convergencia y encontrando soluciones más ralas, y por lo tanto con mayor grado de generalización. Se puede demostrar mediante un sencillo experimento que las mejoras en la clasificación no sólo se deben a la selección de características guiadas por KLD. De hecho, utilizando la información KLD de manera aislada como herramienta de reducción de la dimensionalidad (selección de  $N$  características,  $N = 0,1 \times p$ , asociadas a los  $N$  puntos muestrales con KLD grande), se observó que en promedio la clasificación mediante LDA en las características seleccionadas fue de  $0,55 \pm 0,09$  y  $0,52 \pm 0,03$  para la base de datos P300-1 y P300-2, respectivamente. Estos resultados indican claramente no sólo la importancia de incorporar información a-priori en el modelo, sino también la de atacar la selección de características junto con la clasificación de manera simultánea en un modelo integrativo y guiado por los datos.

Es importante resaltar que ambos métodos fueron implementados utilizando el Algoritmo 1 mediante selección automática de los parámetros de regularización. En vistas del alto grado de rareza en las soluciones acompañado del muy buen desempeño en la clasificación, podemos concluir que el método de selección de los hiperparámetros resulta útil, rápido y eficaz para el tratamiento de estos problemas.

En conclusión, la formulación GSDA resulta una herramienta beneficiosa como método de traducción de señales de EEG.



## Capítulo 5

# Escenario de entrenamiento pequeño: un enfoque para optimizar los tiempos de calibración en BCIs basadas en ERPs

“Nuestro deber es descubrir los privilegios de la simplicidad.”

---

Gabriel García Márquez, Cien años de soledad.

Los resultados que se presentan en este capítulo están orientados a mejorar la detección de ERPs en señales de época única de EEG. Se evalúa la robustez del uso del método GSDA en escenarios adversos con pocos datos de entrenamiento, con el objetivo final de disminuir los tiempos de calibración. Extendiendo estos experimentos así como el método propuesto, al final de este capítulo se presenta un análisis del impacto de la medida de discrepancia utilizada en la construcción de las matrices de anisotropía.

### 5.1. La importancia de disminuir los tiempos de calibración

Tal como se discutió en la Sección [2.4](#) del capítulo introductorio, las BCI típicamente requieren de una gran cantidad de registros de EEG para calibrar (entrenar) correctamente el algoritmo de decodificación. Esto conduce a largas sesiones de entrenamiento donde el usuario se ve sometido a registros de 30 minutos o más. Asimismo las pre-

condiciones fisiológicas del sujeto (horas de sueño, fatiga, concentración, etc.) producen fuertes variaciones en los patrones de la señal de EEG entre sesión y sesión. Por lo tanto es altamente deseable construir algoritmos de traducción que sean capaces de proveer buenos resultados de clasificación ( $\geq 80\%$ ) utilizando la menor cantidad posible de patrones de entrenamiento para así aumentar la practicabilidad de los sistemas BCI al reducir los tiempos de calibración a unos pocos minutos.

La capacidad de generalización de un modelo tiende a aumentar con el número de observaciones de entrenamiento. Al mismo tiempo, a pesar de que usualmente un modelo con más parámetros puede proveer un mejor ajuste a los datos de entrenamiento, la generalización tiende a disminuir con el aumento del número de parámetros en dicho modelo. Por lo tanto, la elección de la cantidad “óptima” de parámetros a incluir requiere de un balance entre la minimización del error de ajuste y la maximización de la capacidad de generalización del modelo (Wolpaw y Wolpaw, 2012).

En esta dirección han surgido varios métodos para la selección eficiente de características en el problema de detección de ERPs inmersos en la señal de EEG. Entre ellos podemos mencionar al método de filtrado espacial basado en el criterio de Fisher (FC, por sus siglas en inglés) (Pires et al., 2011), el cual mejora la SNR y selecciona un número pequeño de características espaciales que luego sirven de entrada a un sencillo clasificador lineal como lo es LDA. Un método bien conocido que realiza selección de características conjuntamente con clasificación es el llamado análisis discriminante lineal paso por paso (SWLDA, por sus siglas en inglés), el cual fue introducido para la clasificación de ERP en (Farwell y Donchin, 1988). Por otro lado, el método de análisis discriminante lineal con acortamiento (SKLDA, por sus siglas en inglés) ha sido introducido en (Blankertz et al., 2011) con el objetivo de mejorar la estimación de las matrices de covarianza en LDA que resultan usualmente mal-condicionadas cuando la cantidad de observaciones es mucho menor que la dimensionalidad de las mismas.

Se pretende entonces estudiar la potencialidad de utilizar GSDA en escenarios de entrenamiento pequeño con el objetivo principal de reducir al mínimo posible la cantidad de registros de EEG requeridos para poder calibrar correctamente el algoritmo de traducción, y, consecuentemente, disminuir los tiempos de calibración requeridos en BCIs basadas en P300. Resultados comparativos entre LDA, SWLDA, SKLDA y FC+LDA son presentados para la base de datos P300-1 y P300-2.

### 5.1.1. Experimentos y resultados

Para estos experimentos se utilizaron las bases de datos P300-1 y P300-2 detalladas en el Capítulo 3. Los escenarios de entrenamiento pequeño se simularon mediante selección aleatoria de un número dado de caracteres a deletrear (2, 4, 6, 8, 10 y 12, para P300-1; 2, 4, 8, 10, 15 y 20 para P300-2). Este proceso de selección se repitió 100 veces. Para disminuir la dimensión de los patrones, las señales se sometieron a submuestreo a 32 y 30 puntos muestrales para la P300-1 y P300-2, respectivamente, produciendo patrones de dimensión 320 ( $32 \times 10$ ) y 448 ( $28 \times 16$ ), respectivamente. Nótese que deletrear *un* caracter genera 180 patrones ( $1 \times 12 \times 15$ ).

En estas simulaciones nuestro método GSDA fue implementado con criterio de parada al seleccionar una cota superior en la cantidad de elementos distintos de cero en el vector solución. En particular para la base de datos P300-1, se comenzó con una cota superior del 40% de la dimensión de los patrones con un incremento del 5% en cada escenario (alcanzando un 65% en el último escenario), mientras que para la base de datos P300-2 estos valores fueron configurados comenzando con un 25% hasta llegar a un 50% en incrementos de un 5%, de la dimensión de los datos. El resto de los métodos, (LDA, SWLDA, SKLDA y FC+LDA) fueron implementados siguiendo los lineamientos dados por los respectivos autores en las correspondientes publicaciones. Para mayores detalles sobre la implementación puede leerse la referencia (Peterson et al., 2017) Sección 5.2 “Small training size scenarios”. Como índice de desempeño aquí también se utilizó AUC. El análisis de significancia estadística entre los resultados de clasificación arrojados por cada método fue realizado mediante ANOVA a dos factores (factor A: escenario, factor B: clasificador) con bloques aleatorizados a un nivel de significancia  $\alpha = 0,05$ . Dado que no se cumple el supuesto de esfericidad (evaluado mediante test de Mauchly,  $p$ -valor  $< 0,05$ ), fue necesario implementar correcciones sobre los grados de libertad y re-ajustar los  $p$ -valores arrojados por ANOVA mediante las correcciones de Greenhouse-Geisser. Encontrando significancia estadística sobre los factores así como sobre su interacción, se procedió luego a evaluar las comparaciones múltiples entre los clasificadores mediante el método de Tukey, también con un nivel de significancia  $\alpha = 0,05$ .

Los resultados comparativos se muestran en la Figura 5.1. Obsérvese que GSDA resulta siempre significativamente mejor ( $p$ -valor  $< 0,05$ ) en comparación con todos los otros métodos para la base de datos P300-1, y para todos los casos excepto para SWLDA para la base de datos P300-2. Dado que la varianza del error entre los métodos es significativamente diferente (evaluada mediante test de Mauchly), se realizó también un ANOVA a dos factores con bloques aleatorizados y corrección de Greenhouse-Geisser sobre las



varianza del error, a un nivel de significancia de  $\alpha = 0,05$ . Se encontró que para la P300-1 la mínima varianza del error siempre corresponde a GSDA ( $p$ -valor  $< 0,05$ ), tal como puede observarse en la Figura 5.1. Estos resultados muestran que GSDA constituye un clasificador más robusto, y por lo tanto más estable y confiable.

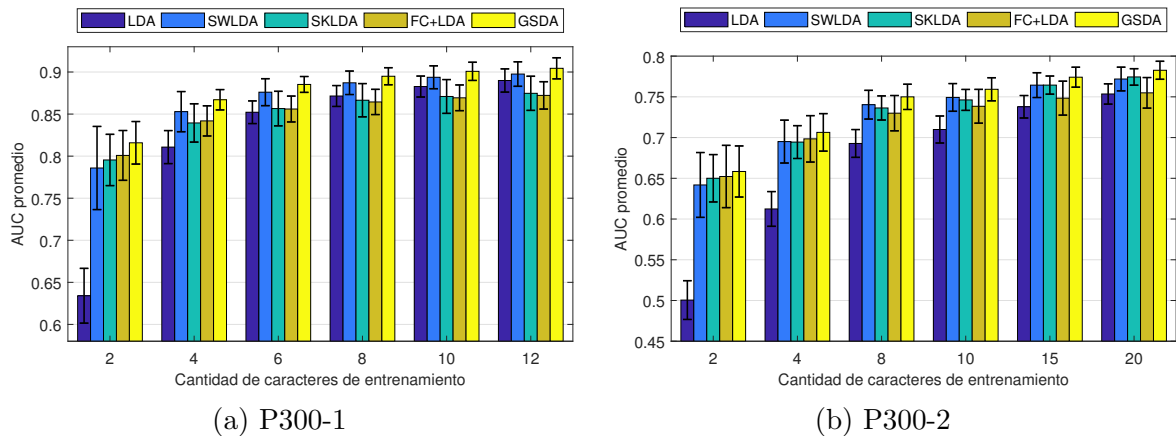


Figura 5.1: Promedio de AUC para los datos de testeo alcanzados por LDA, SWLDA, SKLDA, FC+LDA y GSDA, respectivamente, en cada uno de los escenarios de entrenamiento pequeño simulado para las bases de datos P300-1 y P300-2. La barra de error indica la desviación estándar de AUC a lo largo de las 100 repeticiones.

A la luz de estos resultados es razonable considerar a GSDA como un buen candidato para disminuir los tiempos de calibración en sistemas BCI basados en ERPs. Si se decide calibrar el sistema con los patrones provenientes de deletrear sólo 10 caracteres, se estima que es necesario contar con alrededor de 5 min para la toma de los datos (se asume 100 ms de tiempo de intensificación del estímulo, 75 ms de tiempo inter-estímulo, 12 estimulaciones por bloque y 15 repeticiones de cada bloque de estimulación). En nuestros experimentos encontramos que para entrenar completamente GSDA en el escenario “10”, es decir utilizando 1800 patrones de entrenamiento, se requieren de sólo 5 a 10 s. Este insignificante costo computacional sumado a los excelentes valores de clasificación alcanzados ( $> 90\%$  para P300-1 y  $> 75\%$  para P300-2), hacen de GSDA un muy buen candidato para la detección en tiempo real de P300 inmersos en la señal de EEG.

## 5.2. Medidas de discrepancia: un análisis de su impacto en la detección de ERPs

Como se ha demostrado en el Capítulo 4, la inclusión de información discriminativa a-priori en el problema de regularización con restricciones no sólo mejora la precisión, sino que también acelera la convergencia del clasificador, en comparación a cuando dicha información no es considerada. Si bien GSDA ha sido sólo evaluado utilizando una versión simétrica de la divergencia de Kullback-Leibler (ec. 4.13), su formulación admite la utilización de cualquier medida que cuantifique la discrepancia entre clases. Por lo tanto, resulta natural explorar el impacto de la utilización de diferentes medidas de discrepancia para resaltar la información del P300 contenida en los datos de entrenamiento.

Dentro de la familia de entropías de Shannon, además de la divergencia de Kullback-Leibler (4.12) y su versión simétrica (4.13), podemos considerar la divergencia de Jensen-Shannon (JS) (Lin, 1991). Sean  $f_1(n)$  y  $f_2(n)$ ,  $n \in \mathcal{N}$ , como antes, dos funciones de probabilidad correspondientes a las variables aleatorias discretas  $\mathbf{W}_1$  y  $\mathbf{W}_2$ , respectivamente, en una cierta característica  $i$ . Se define entonces la divergencia-JS,  $d_{JS}$ , como:

$$d_{JS_\pi}(f_1, f_2) \doteq H(\pi f_1 + (1 - \pi)f_2) - \pi H(f_1) - (1 - \pi)H(f_2), \quad (5.1)$$

donde  $H(\cdot)$  es la entropía de Shannon  $H(f) \doteq -\sum_{n \in \mathcal{N}} f(n) \log(f(n))$  (Shannon, 2001) y  $\pi$ , es un parámetro predefinido tal que  $0 \leq \pi \leq 1$ . En el caso particular  $\pi = 1/2$ , la divergencia JS es simétrica y su raíz cuadrada es una métrica en el estricto sentido matemático (Lin, 1991).

La distancia euclídea en  $\mathbb{R}$  también puede utilizarse para definir una medida de discrepancia denominada distancia total variacional, definida por la norma  $\ell_1$ :

$$d_V(f_1, f_2) \doteq \sum_{n \in \mathcal{N}} |f_1(n) - f_2(n)| = \|f_1 - f_2\|_1. \quad (5.2)$$

Similarmente, si utilizamos la norma  $\ell_2$ , obtenemos la conocida distancia Euclídea:

$$d_E(f_1, f_2) \doteq \sqrt{\sum_{n \in \mathcal{N}} (f_1(n) - f_2(n))^2} = \|f_1 - f_2\|_2. \quad (5.3)$$

Otra medida de discrepancia muy utilizada en el contexto de aprendizaje maquina, es la distancia de Fisher. Consideremos el caso en que  $f_1$  y  $f_2$  son funciones de densidad de probabilidad de dos distribuciones Gaussianas con parámetros  $\mu_1, \sigma_1$  y  $\mu_2, \sigma_2$ ,

respectivamente. La distancia de Fisher entre  $f_1$  y  $f_2$  puede calcularse como (Costa et al., 2015):

$$d_F(f_1, f_2) = \sqrt{2} \ln \left( \frac{\mathcal{F}((\mu_1, \sigma_1), (\mu_2, \sigma_2)) + (\mu_1 - \mu_2)^2 + 2(\sigma_1^2 + \sigma_2^2)}{4\sigma_1\sigma_2} \right), \quad (5.4)$$

donde  $\mathcal{F}((\mu_1, \sigma_1), (\mu_2, \sigma_2)) = \sqrt{((\mu_1 - \mu_2)^2 + 2(\sigma_1 - \sigma_2)^2)((\mu_1 - \mu_2)^2 + 2(\sigma_1 + \sigma_2)^2)}$ .

Finalmente, la información discriminativa puede ser también evaluada mediante el coeficiente de correlación al cuadrado con signo, denominado “valor  $r^2$  con signo”. Comenzamos definiendo  $r$ :

$$r = \frac{\sqrt{q}}{q+1} \left( \frac{\mu_1 - \mu_2}{\sigma} \right), \quad (5.5)$$

donde  $q$  es la proporción de elementos de la clase 1 relativa a los elementos de la clase 2,  $\sigma$  es la desviación estándar conjunta de ambas poblaciones y  $\mu_1$  y  $\mu_2$  son sus respectivas medias. Luego, el valor  $r^2$  con signo se define como:

$$d_r(f_1, f_2) = \text{sgn}(r) r^2. \quad (5.6)$$

Estas medidas, tal como la divergencia-J (4.13), pueden ser utilizadas en la construcción de las matrices de anisotropía  $\mathbf{D}_1$  y  $\mathbf{D}_2$ . Para probar la relevancia del uso de pesos adecuados en la penalización, las diferentes medidas de discrepancia se comparan contra una medida de “ignorancia”, denotada mediante  $d_{rnd}$ , la cual simplemente pesa cada punto muestral (característica) en función de un número aleatorio entre 0 y 1.

### 5.2.1. Experimentos y resultados

Para analizar el impacto de la cantidad de patrones utilizados para estimar las distribuciones de probabilidad, se simularon escenarios de entrenamiento pequeño para el problema de detección de ERPs inmersos en la señal de EEG, tal como se realizó en 5.1.1. Cada medida de discrepancia considerada ( $d_J$ ,  $d_{JS_{\pi_1, \pi_2}}$ ,  $d_V$ ,  $d_E$ ,  $d_F$  y  $d_r$ ), como así la medida de ignorancia  $d_{rnd}$  fue utilizada para construir las matrices de anisotropía  $\mathbf{D}_1$  y  $\mathbf{D}_2$ . Las distribuciones discretas de probabilidad  $f_1^i$  y  $f_2^i$ ,  $i = 1, \dots, p$ , fueron estimadas mediante histogramas. Para ilustrar el efecto del peso  $\pi$  en la divergencia JS se analizaron dos casos: i)  $d_{JS_{1/2}}$ , pesos iguales a un medio,  $\pi = \frac{1}{2}$  y ii)  $d_{JS_{\pi}}$ , pesos basados en las prior de cada clase  $\pi = \pi_k = \frac{n_k}{N}$ . Note que con ésta última elección, la divergencia JS es equivalente a la información mutua (Grosse et al., 2002). Por último, dado que  $d_r$  puede tomar valores negativos, (Ec. (5.6)), para su uso en la construcción de las matrices  $\mathbf{D}_1$  y  $\mathbf{D}_2$ , se

transformaron sus valores linealmente entre 0 y 1.

### Gráficas discriminativas

Las señales de ERP son señales localizadas en el tiempo, que pueden ser fácilmente estimadas mediante la promediación coherente de un número de patrones provenientes de la clase de interés. Cada una de las siete medidas de discrepancia considerada evalúa la distancia entre clases en cada punto del espacio tiempo-canal (pixel, punto muestral). Las Figuras 5.2a, 5.2b, y 5.3a, 5.3b muestran los grandes promedios de las clases con y sin P300, y las gráficas en el espacio tiempo-canal correspondientes a cada medida considerada en dos escenarios de entrenamiento extremos: a)  $n \lesssim p$  (Escenario-1), y b)  $n \gg p$  (Escenario-6), respectivamente. Nótese que, en concordancia con los grandes promedios para la clase con P300, todas las medidas resaltan un área de características relevantes en una ventana alrededor de los 300 ms. Si se compara la dispersión de estas áreas entre los dos escenarios considerados, se observa claramente que dicha región se torna más prominente a medida que la cantidad de observaciones aumenta, dado que las estimaciones de las distribuciones de probabilidad mejora a medida que aumenta el número de muestras (véanse las Figuras 5.2b y 5.3b). Por otro lado, es también interesante destacar que las medidas que asumen normalidad en los datos y se basan sólo en los dos primeros momentos estadísticos, parecen descartar toda información que no provenga de la prominencia de la onda P300. Por último, nótese cómo aquí también se observa que para la base de datos P300-2, las contribuciones de ciertos canales a la discrepancia entre clases parecen ser despreciables.

### Impacto en la clasificación

Las Figuras 5.4a y 5.4b muestran el desempeño general en términos de AUC de cada medida para cada escenario de entrenamiento para las bases de datos P300-1 y P300-2, respectivamente. En primer lugar, estos resultados muestran que el uso de cualquier medida de discrepancia siempre resulta mejor ( $p$ -valor  $< 0,01$  para P300-1) que cuando se utilizan los pesos aleatorios ( $d_{rnd}$ ). Segundo, nótese como las diferencias entre los desempeños de las medidas son visualmente distinguibles para los escenarios de escasos datos de entrenamiento (deletreo de 2 y 4 caracteres), llegando a resultados similares a medida que aumenta el número de patrones de entrenamiento. Probablemente esto se deba a que a medida que aumenta la cantidad de patrones de entrenamiento, las estimaciones de las distribuciones muestrales  $f_1$  y  $f_2$  se vuelven más confiables, y por lo tanto todas las me-

didadas resultan igualmente buenas. Finalmente, es interesante notar que aquellas medidas que aparentemente mejor resaltaban la onda P300 en el espacio tiempo-canal ( $d_F$  y  $d_r$ , Figuras 5.2 y 5.3) resultan siempre peores en términos de clasificación. Esto puede deberse a que ambas medidas asumen normalidad en los datos, y por lo tanto las diferencias entre clases son medidas en función de la distancia entre las medias y entre las varianzas de las distribuciones.

El análisis estadístico muestra que no existen diferencias significativas entre los resultados arrojados por las medidas provenientes de la familia de Shannon y las basadas en la norma  $p$ ,  $p = 1, 2$ , ( $d_J$ ,  $d_{JS_{1/2}}$ ,  $d_{JS_\pi}$ ,  $d_V$ ,  $d_E$ ). Asimismo, para la base de datos P300-1 se encontró que estas medidas resultan siempre mejores ( $p$ -valor  $< 0,01$ ) que aquellas que asumen normalidad ( $d_F$ ,  $d_r$ ). Si bien para la base de datos P300-2 estas diferencias no son significativas, los resultados también muestran claramente que siempre resulta mejor, en términos de clasificación, trabajar sin ninguna suposición sobre la distribución de los datos.

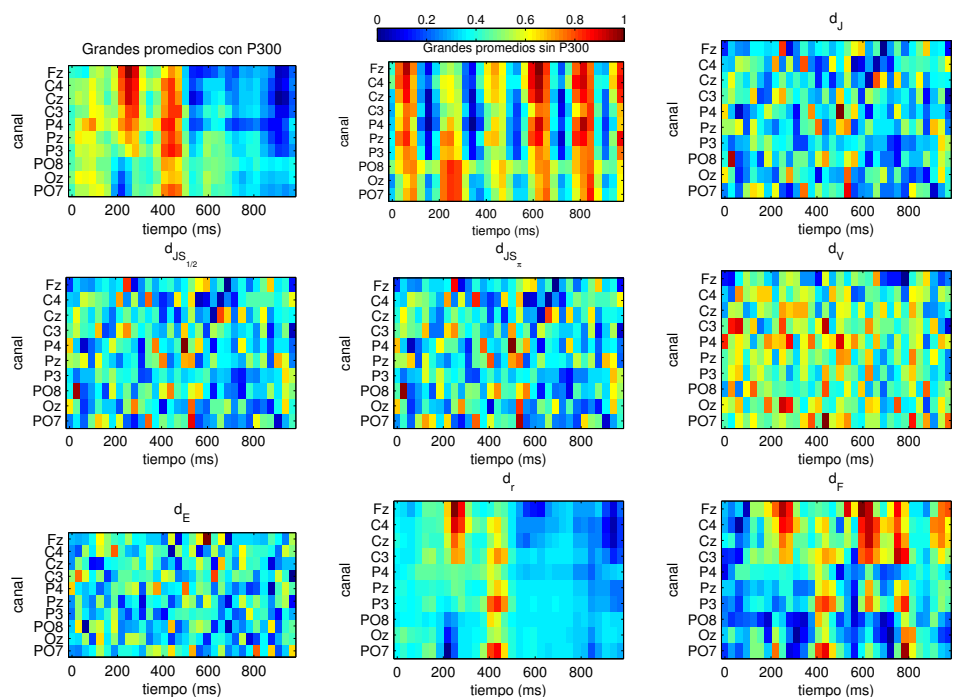
Estos resultados extienden las investigaciones ya analizados en la sección anterior 5.1.1, Figura 5.1 (ver también (Peterson et al., 2017)), en el cual GSDA fue implementado con la medida  $d_J$ . El análisis del impacto de las medidas de discrepancia utilizadas para construir las matrices de anisotropía utilizadas en GSDA nos permite concluir que medidas basadas en la divergencia de Kullback-Leibler con ninguna suposición sobre la distribución de las clases es una buena elección para cuantificar la distancia entre señales de EEG con ERP y aquellas sin ERP.

### 5.3. Discusiones

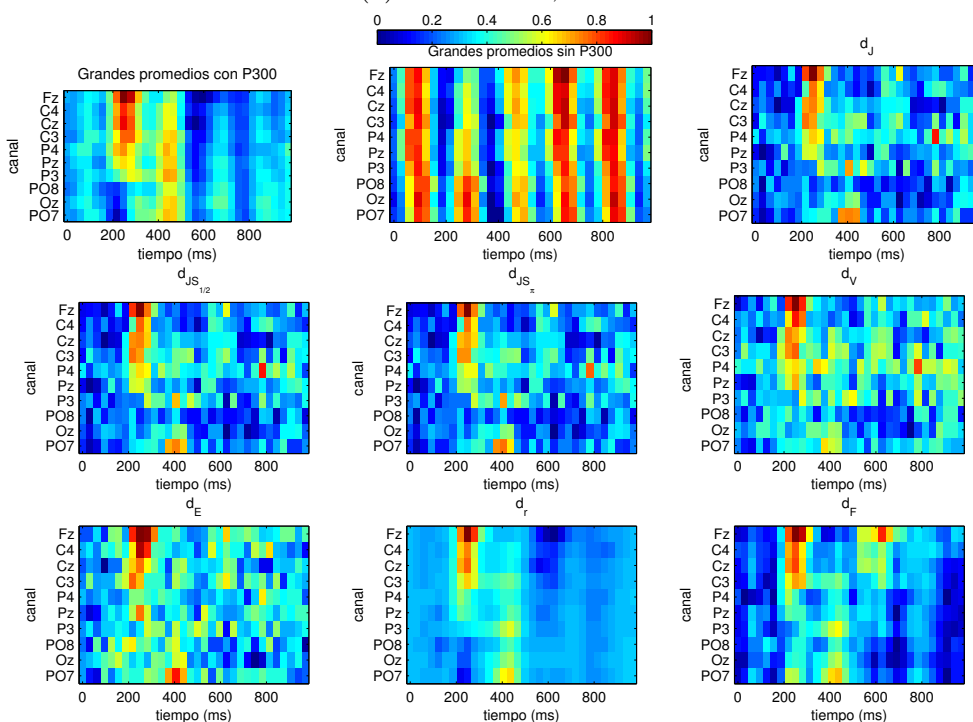
En este capítulo hemos presentado el desempeño de GSDA en escenarios de entrenamiento pequeño con el objetivo de disminuir los tiempos de calibración de los sistemas BCI basados en ERPs. El desempeño superior de GSDA comparado con el obtenido mediante las otras versiones regularizadas de LDA, muestran no sólo la importancia de atacar la selección de características junto con la clasificación, sino también como la apropiada inclusión de información discriminativa a-priori es capaz de mejorar el desempeño del clasificador. Asimismo, a la luz de la alta tasa de clasificación junto con la baja varianza encontrada a lo largo de las 100 iteraciones de entrenamiento (por ejemplo, Escenario-10, 5.1), vemos que GSDA resulta un excelente candidato para calibrar diariamente una BCI de manera confiable en pocos segundos. De este modo se espera que el sistema sea capaz de adaptarse mejor a las posibles variaciones en las pre-condiciones fisiológicas del sujeto,

con el menor tiempo requerido.

El análisis del impacto de las diferentes medidas de discrepancia confirman nuestra hipótesis de que la inclusión apropiada de información a-priori tiene un impacto directo en la clasificación. Los resultados comparativos de GSDA con diferentes medidas de discrepancia mostraron que aquellas medidas basadas en los primeros dos momentos estadísticos resultan siempre peores en términos de clasificación, mientras que para el resto de las medidas no hay diferencias significativas en el uso de las mismas. De todas maneras, confrontados con la necesidad de seleccionar solo una de ellas, la divergencia JS no sólo presenta propiedades matemáticas interesantes, sino que también permite extender fácilmente su formulación a problemas multiclase.

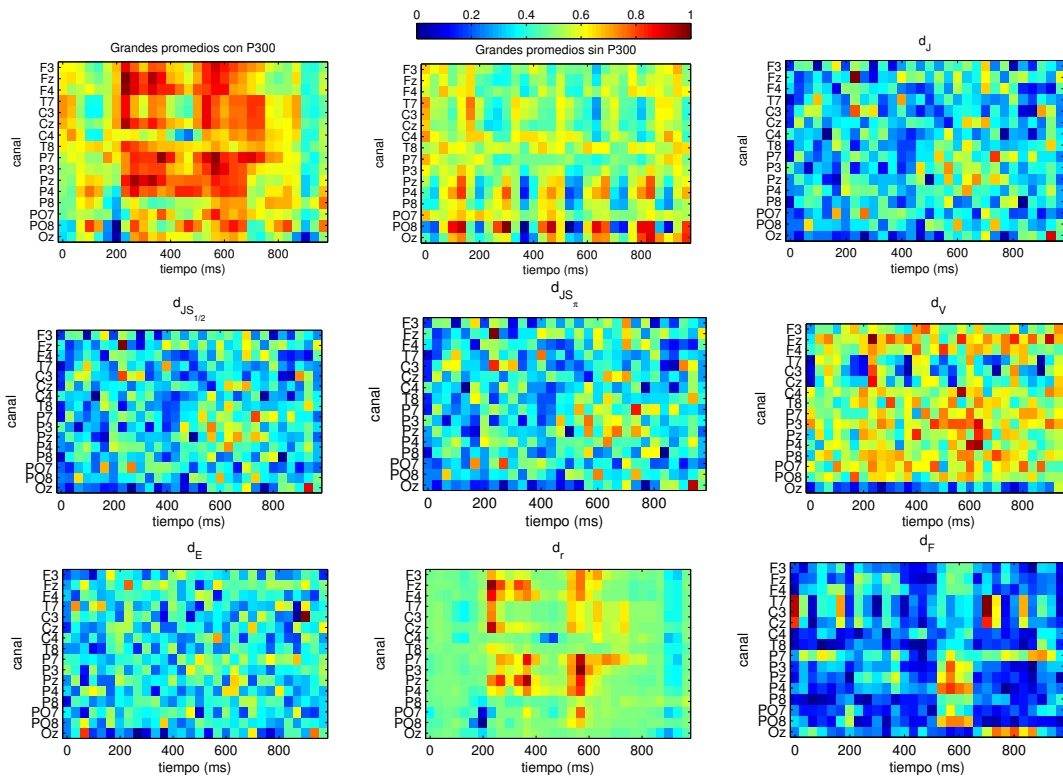


(a) Escenario-1, P300-1

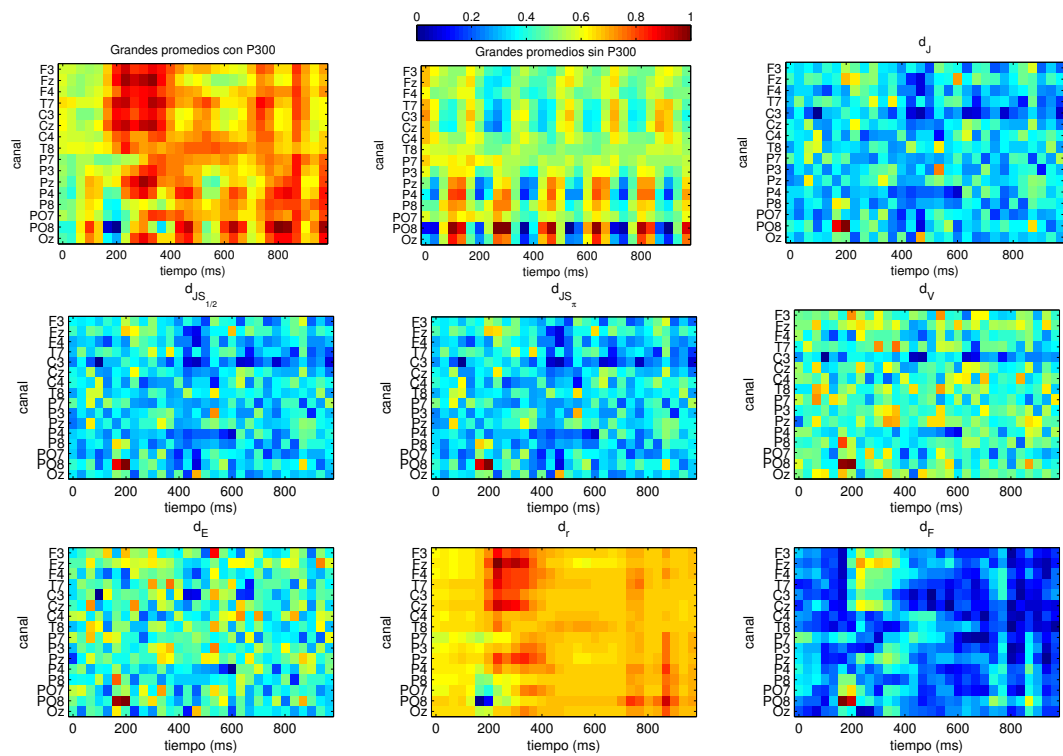


(b) Escenario-6, P300-1

Figura 5.2: Grandes promedios y gráficas discriminativas para un sujeto de la base de datos P300-1. De izquierda a derecha y de arriba a abajo se presenta en el plano tiempo-canal los mapas discriminativos calculados mediante: promediación coherente para la clase con P300, promediación coherente para clase sin P300,  $d_J$ ,  $d_{J_{S_{\pi_1, \pi_2}}}$ ,  $d_V$ ,  $d_E$ ,  $d_F$  y  $d_r$ . a) Escenario-1 b) Escenario-6



(a) Escenario-1



(b) Escenario-6

Figura 5.3: Grandes promedios y gráficas discriminativas para un sujeto de la base de datos P300-2. De izquierda a derecha y de arriba a abajo se presenta en el plano tiempo-canal los mapas discriminativos calculados mediante: promediación para la clase con P300, promediación coherente para clase sin P300,  $d_J$ ,  $d_{JS_{\pi_1, \pi_2}}$ ,  $d_V$ ,  $d_E$ ,  $d_F$  y  $d_r$ . a) Escenario-1 b) Escenario-6



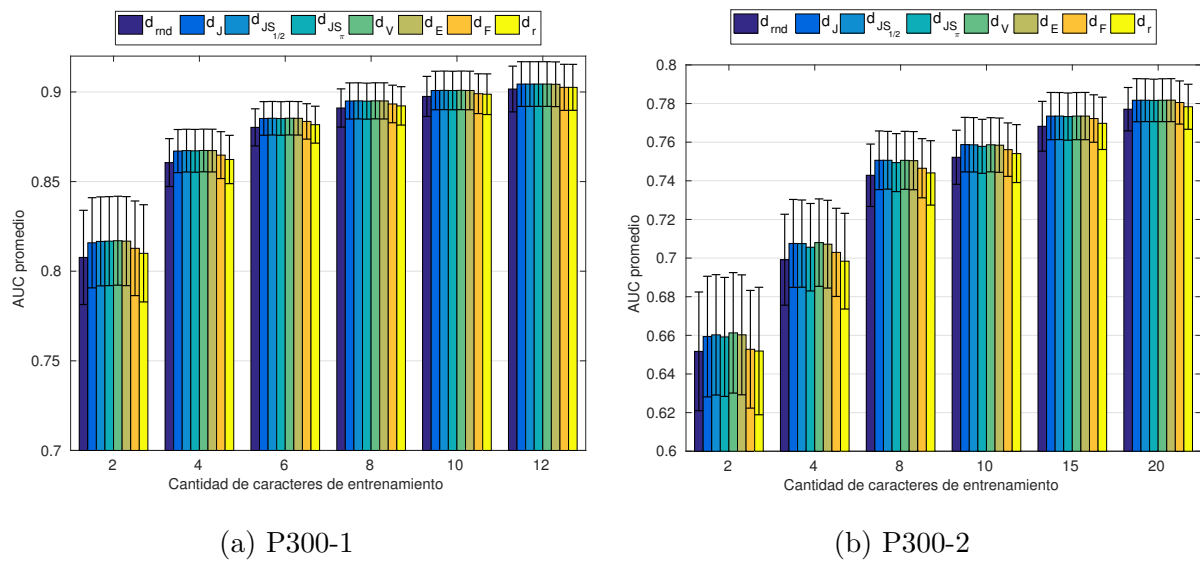


Figura 5.4: Resultados de clasificación promedio en términos de AUC en cada escenario de entrenamiento simulado al utilizar las diferentes medidas de discrepancia consideradas dentro de GSDA.

## Capítulo 6

# Un método integrador para la detección de la intención de movimiento

“Tenía la idea de que al poner nombre a los problemas, éstos se materializan y ya no es posible ignorarlos; en cambio, si se mantienen en el limbo de las palabras no dichas, pueden desaparecer solos, con el transcurso del tiempo.”

---

Isabel Allende, La casa de los Espíritus.

En capítulos anteriores nos hemos centrado en atacar la detección de ERPs inmersos en la señal de EEG. En este capítulo, haciendo uso del método GSDA, nos centraremos en la construcción de algoritmos para mejorar la detección de la intención de movimiento aplicado a terapias de rehabilitación basadas en BCIs de pacientes con severa pérdida del control motor de alguna de sus extremidades superiores.

### 6.1. Filtrado espacial y banco de filtros

Los patrones asociados a imaginar la realización de ciertos movimientos, tal como se mencionó en la Sección [2.2.2](#), pueden distinguirse espacialmente a lo largo de la corteza sensoriomotora. Resulta entonces natural explorar técnicas de filtrado espacial que permitan distinguir entre dos tipos de condiciones de MI (por ejemplo, MI de la mano izquierda vs. MI de la mano derecha). El método denominado “patrones espaciales comunes” (CSP, por sus siglas en inglés), maximiza la varianza de una señal de EEG filtrada para una

condición (clase) mientras que la minimiza con respecto a la otra (Ramoser et al., 2000). Además, dado que la varianza de una señal filtrada es igual a su banda de potencia, el método CSP provee un enfoque adecuado para distinguir entre dos condiciones de MI (Blankertz et al., 2008). Como desventaja, si bien utilizando CSP es posible alcanzar porcentajes de clasificación por encima del 75%, su éxito depende en gran medida de la banda de filtrado pre-definida (típicamente de 8-30 Hz) y de la longitud del segmento de EEG extraído luego del comienzo de la instrucción de MI (generalmente de 0,5-2,5 s) (Blankertz et al., 2008). Una manera, intuitiva aunque poco práctica, de seleccionar la banda frecuencial óptima para cada sujeto es mediante búsqueda exhaustiva y selección manual. Si bien se ha probado que mediante esta metodología se puede mejorar la detección de MI, es una técnica no estandarizada, computacionalmente costosa y difícil de implementar (Novi et al., 2007). El método CSP y banco de filtros (FBCSP, por sus siglas en inglés), propuesto por Ang et al. (Ang et al., 2008) para la selección automática de bandas frecuenciales, es un método de cuatro pasos en el cual CSP es aplicado a diferentes bandas frecuenciales prefijadas, y donde las características más relevantes son seleccionadas de a pares según un criterio basado en la información mutua. Una versión mejorada de FBCSP, denominada FBCSP discriminativo (DFBCSP, por sus siglas en inglés) (Thomas et al., 2009), arroja mejores tasas de clasificación al realizar banco de filtros para cada sujeto basado en el criterio de Fisher previo a la extracción de características. Más recientemente, se propuso el método FBCSP ralo (SFBCSP, por sus siglas en inglés) (Zhang et al., 2015b), en el cual un número pequeño de características CSP sub-banda son automáticamente seleccionadas mediante el método LASSO. El método SFBCSP ha mostrado mejores resultados de clasificación en comparación con los métodos CSP, FBCSP y DFBCSP en bases de datos MI-EEG públicamente accesibles (Zhang et al., 2015b). Como desventaja, SFBCSP utiliza validación cruzada para la estimación del parámetro de regularización y no tiene en consideración información previa alguna sobre la discrepancia entre clases.

Además de seleccionar las bandas frecuenciales que mejor representan los patrones ERD/ERS en cuestión, varios grupos han estudiado el efecto de la longitud y de la latencia del segmento de EEG extraído respecto al comienzo de la instrucción de MI. En particular, Ang et al. (Ang et al., 2012b) extendieron su método FBCSP mediante la aplicación de bancos de filtros en tres segmentos de EEG prefijados, en el cual el segmento temporal óptimo es seleccionado como aquel de máxima información mutua en un proceso de validación cruzada. Por otro lado, recientemente fue presentado el método CSP segmento tiempo-frecuencia ralo (STFSCSP, por sus siglas en inglés) (Miao et al., 2017). Este

método, luego de realizar una apropiada selección de canales para cada sujeto, mejora al método SFBCSP al dividir cada banda frecuencial en diferentes ventanas temporales.

En este capítulo presentamos un método que busca superar las principales limitaciones del uso de segmentos temporo-frecuenciales prefijados en BCIs basadas en MI. En lo subsiguiente, vamos a denominar a éste método CSP y bandas temporo-frecuenciales penalizadas (PTFBCSP, por sus siglas en inglés), en el cuál las bandas espectro-temporales óptimas para cada sujeto serán automáticamente seleccionadas. Las principales características de este método son: i) permite analizar la señal de EEG en múltiples ventanas de tiempo simultáneamente; ii) en cada ventana temporal se realiza banco de filtros; iii) las características espacio-espectro-temporales extraídas mediante CSP son consideradas independientes; y iv) la selección y la clasificación de las características CSP espectro-temporales se realizan simultáneamente mediante GSDA.

A continuación presentamos una breve descripción del método así como una introducción a CSP para MI-BCIs.

## 6.2. Patrones espaciales comunes y bandas temporo-frecuenciales penalizadas

El método propuesto, que se encuentra esquemáticamente representado en la Figura [6.1](#), es una combinación de algoritmos de procesamiento de señales y de aprendizaje maquina. En un primer paso, cada señal de EEG multicanal cruda se descompone en  $T$  ventanas temporales ( $tb_1, \dots, tb_t, \dots, tb_T$ ). En segundo lugar, cada segmento  $t$ -ésimo se divide en  $F$  bandas frecuenciales ( $fb_1, \dots, fb_f, \dots, fb_F$ ). Luego, las características espaciales en cada banda temporo-frecuencial son extraídas mediante CSP. La selección de características junto con la clasificación es automáticamente realizada por GSDA.

En la Figura [6.1](#) vemos que PTFBCSP está esencialmente compuesto por dos algoritmos: CSP para la extracción de características y GSDA para la selección y clasificación de características. En lo que sigue, haremos un breve resumen del método CSP. Para más detalles sobre GSDA, se remite al lector al Capítulo 4.

### Extracción de características mediante CSP

El método CSP es uno de los algoritmos más populares para la extracción de características en MI-BCIs. Esto se debe a que dadas dos épocas de EEG registradas bajo dos condiciones de MI diferentes, CSP busca una transformación lineal tal que en el espacio

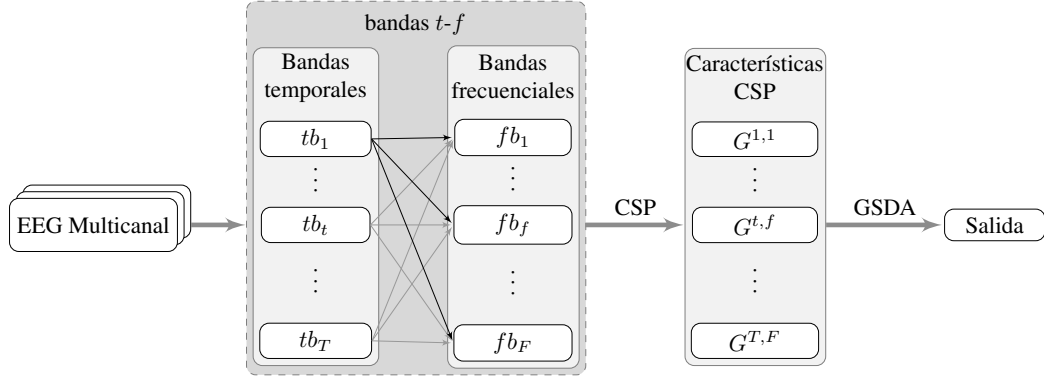


Figura 6.1: Representación esquemática del método PTFBCSP propuesto. La señal de EEG multicanal es analizada en  $T$  ventanas temporales. Luego, cada  $t$ -ésimo segmento temporal es dividido en  $F$  bandas frecuenciales. El método CSP es aplicado a cada segmento temporo-frecuencial. Finalmente, la selección de las características CSP temporo-frecuenciales óptimas es automáticamente realizada junto con la clasificación mediante GSDA.

transformado la varianza sea máxima para una clase y mínima para la otra. Este método fue utilizado por primera vez en (Ramoser et al., 2000) para distinguir entre MI de la mano derecha y MI de la mano izquierda en señales de EEG multicanal.

Sea  $\{\mathbf{X}_c^i\}_{i=1}^{n_c}$  un conjunto de  $n_c$  señales de EEG filtradas pasa-banda,  $c = 1, 2$ . Aquí, cada  $\mathbf{X}_c^i$  es una matriz de  $p \times m$ , donde  $p$  y  $m$  denotan el número de canales y el número de puntos muestrales en cada canal, respectivamente. En lo que sigue, vamos a suponer que cada época  $\mathbf{X}_c^i$  está normalizada, es decir tiene media cero y varianza unitaria. El objetivo de CSP es encontrar  $p$  filtros espaciales,  $\mathbf{w}_1, \dots, \mathbf{w}_p \in \mathbb{R}^p$ , para linealmente transformar la señal de entrada  $\mathbf{X}$  mediante:

$$\mathbf{Z} = \mathbf{W}' \mathbf{X}, \quad (6.1)$$

donde  $\mathbf{W} = [\mathbf{w}_1, \dots, \mathbf{w}_p]$  y  $'$  denota la traspuesta.

El criterio de optimización para hallar los  $p$  filtros espaciales equivale a resolver el siguiente problema de autovalores generalizado<sup>1</sup>:

$$\Sigma_1 \mathbf{w}_j = \gamma_j \Sigma_2 \mathbf{w}_j, \quad j = 1, \dots, p, \quad (6.2)$$

en donde

$$\Sigma_c = \frac{1}{n_c} \sum_{i=1}^{n_c} \mathbf{X}_c^i (\mathbf{X}_c^i)', \quad c = 1, 2.$$

<sup>1</sup>Detalles del procedimiento matemático pueden encontrarse en el Anexo A.2 de esta tesis.

Los autovalores generalizados  $\gamma_j$  cuantifican la relación de varianzas entre clases en la característica  $j$ . Un valor de  $\gamma_j$  grande indica que el correspondiente filtro espacial  $\mathbf{w}_j$  contiene valores de varianza grandes para la clase 1 pero bien bajos para la clase 2, mientras que para un valor de  $\gamma_j$  pequeño ocurre lo opuesto. Además, nótese que  $\mathbf{W}$  es no-singular dado que sus vectores columna  $\mathbf{w}_j$  son linealmente independientes. Por lo tanto, de la ec. (6.1) obtenemos  $\mathbf{X} = (\mathbf{W}^{-1})' \mathbf{Z}$ . Definiendo  $\mathbf{A} = (\mathbf{W}^{-1})' \in \mathbb{R}^{p \times p}$ , se denominan “patrones espaciales” a los correspondientes vectores columna  $\mathbf{a}_j \in \mathbb{R}^p$ ,  $j = 1, \dots, p$ , de la matriz  $\mathbf{A}$ .

En general, solo un número pequeño  $K$  de filtros espaciales es utilizados para computar la extracción de características. Luego de re-ordenar los filtros espaciales en orden descendente en función de  $\gamma_j$ , definimos la matriz  $\mathbf{Z}_K^*$  de  $2K \times p$ , resultado de apilar a la derecha las  $K$  primeras y últimas filas de  $\mathbf{Z}$ , como sigue  $\mathbf{Z}_K^* = [\mathbf{z}_1, \dots, \mathbf{z}_K, \mathbf{z}_{p-(K+1)}, \dots, \mathbf{z}_p]'$ . Luego, el vector de características CSP se define como  $\mathbf{g} \doteq [g_1, \dots, g_{2K}]'$ , donde

$$g_j \doteq \log(S^2(\mathbf{z}_j^*)), \quad j = 1, \dots, 2K, \quad (6.3)$$

con  $S^2(\mathbf{z}_j^*) \doteq \frac{1}{p} \sum_{\ell=1}^p (\mathbf{z}_j^{*\ell} - \mu_j)^2$ ,  $\mu_j \doteq \frac{1}{p} \sum_{\ell=1}^p \mathbf{z}_j^{*\ell}$ .

En el contexto de PTFBCSP, el algoritmo CSP es implementado en cada banda frecuencial  $f$  perteneciente a un segmento de tiempo particular  $t$ . Esto resulta en  $2KF$  características extraídas por cada banda tiempo-frecuencia por cada época de EEG multicanal. Mediante la concatenación de estas bandas características CSP temporo-frecuenciales y apilando los vectores resultantes correspondientes a cada una de las  $n$  muestras de entrenamiento, construimos la siguiente matriz de características  $\mathbf{G}$  de dimensión  $n \times q$ ,  $q=2KFT$ :

$$\mathbf{G} \doteq \begin{bmatrix} g_{1,1} & \dots & g_{1,2KF} & \dots & g_{1,2KFT} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ g_{n,1} & \dots & g_{n,2KF} & \dots & g_{n,2KFT} \end{bmatrix}, \quad (6.4)$$

donde  $g_{i,j}$  denota la  $j$ -ésima característica tiempo-frecuencia extraída de la  $i$ -ésima época de EEG.

## Selección de características y clasificación via GSDA

Seleccionar las características más relevantes de la matriz  $\mathbf{G}$  equivale a detectar las bandas frecuenciales “óptimas” a lo largo de los segmentos temporales. La interacción entre estas características debe tenerse en consideración para analizar apropiadamente su

contribución a la discriminación de clases. En este sentido, a sabiendas de las propiedades de discriminabilidad de GSDA (Peterson et al., 2017), hemos utilizado dicho método para la selección automática de características junto con la clasificación.

En este contexto el grado de discriminabilidad se calcula para cada elemento  $g_{i,j}$  de la matriz  $\mathbf{G}$ , suponiendo que dicho elemento es una realización de alguna de las dos variables aleatorias cuyas distribuciones unívocamente caracterizan a cada una de las dos condiciones de MI. Basándonos en los resultados hallados en el Capítulo 5, proponemos cuantificar dicho poder de discriminabilidad mediante la divergencia de Jensen-Shannon (Lin, 1991).

El método GSDA, presentado detalle la Sección 4.2, busca hallar el vector discriminativo óptimo  $\hat{\beta}$ , que contenga unos pocos valores distintos de cero que corresponden a las bandas  $t$ - $f$  más relevantes. Tal como antes, una regla de clasificación sencilla puede entonces construirse sobre la matriz de datos proyectados  $\mathbf{G}^* \doteq \mathbf{G}\hat{\beta}$  de dimensión  $n \times 1$ .

### 6.3. Experimentos y resultados

En este capítulo presentamos los resultados experimentales utilizando las bases de datos MI-1, MI-2 Y MI-3, detalladas en el Capítulo 3, Subsecciones 3.2.1, 3.2.2 y 3.2.3. Dos configuraciones diferentes de PTFBCSP fueron evaluadas, utilizando múltiples ventanas temporales y también el caso particular con  $T = 1$ , es decir, utilizando una única ventana temporal. A esta última configuración, en lo que sigue la denominaremos como CSP y bandas frecuenciales penalizadas (PFBCSP, por sus siglas en inglés). Estas dos configuraciones fueron elegidas para analizar las siguientes dos conjeturas: i) la introducción de información discriminativa a-priori para la selección de bandas en un modelo integrado es beneficiosa para realzar la separabilidad entre dos condiciones de MI, y ii) el uso combinado de características (información) temporo-frecuenciales puede producir una clasificación no solo más precisa sino también más robusta. El análisis comparativo de la eficiencia del método propuesto en sus dos configuraciones para detectar MI, fue realizado contra el tradicional CSP y el reciente mejor método SFBCSP, ya que este último ha mostrado mejores resultados de clasificación en comparación con los métodos DFBCSP y FBCSP (Zhang et al., 2015b).

### 6.3.1. Análisis de los datos

Con el objetivo de reducir la dimensionalidad de las señales, y para potencialmente disminuir los tiempos de preparación del sujeto en aplicaciones en tiempo real, para las bases de datos MI-1 y MI-2, se seleccionaron 28 electrodos cubriendo las regiones sensoriomotoras, tal como se indica en (Marchesotti et al., 2017). De esta manera, se evita el uso de señales cuya información no está relacionada con la intención del movimiento.

Con el objetivo de analizar separadamente el impacto de usar, por un lado, GSDA como método integrativo para la selección y clasificación, y por otro lado, las múltiples ventanas temporales en el análisis de la señal de EEG, nuestro método PTFBCSP fue primero evaluado en su versión de ventana temporal única. En esta configuración, a la que llamamos PFBCSP, para cada época de EEG se extrajo un único segmento temporal entre 0,5 y 2,5 s, luego del inicio de la señal de MI visual. Para el algoritmo CSP, cada segmento de EEG fue filtrado mediante un filtro Chebyshev tipo II entre 4-40 Hz, mientras que para los algoritmos de banco de filtros (SFBCSP y PFBCSP), 17 sub-bandas ( $F = 17$ ) entre 4 y 40 Hz con 4 Hz de ancho de banda y 2 Hz de solapamiento fueron construidas tras aplicar un filtro Chebyshev tipo II ( $fb_1 = 4 - 8$  Hz,  $fb_2 = 6 - 10$  Hz,  $\dots$ ,  $fb_F = 36 - 40$  Hz). Este procedimiento sigue la metodología utilizada en (Thomas et al., 2009; Zhang et al., 2015b). El número final de pares de filtros CSP fue fijado en 1 ( $K = 1$ ). Por lo tanto, luego de aplicar CSP a cada sub-banda, se construyó un vector de características de dimensión  $q = 2KF = 34$  para cada segmento de EEG. La dimensión de la matriz  $\mathbf{G}$  (Ec. 6.4) es entonces de  $n \times 34$ , donde  $n$  corresponde al número de patrones de entrenamiento. Finalmente, a modo de evitar estimaciones pobres de las matrices de covarianza (lo cual generalmente ocurre cuando se cuenta con pocos datos de entrenamiento), el algoritmo CSP fue implementado mediante contracción automática de las matrices de covarianza estimadas, tal como se propone en (Lotte, 2015).

El método SFBCSP fue implementado como lo describen sus autores en (Zhang et al., 2015b). Por lo tanto, el problema de regresión LASSO fue implementado utilizando el algoritmo de gradiente descendiente (Friedman et al., 2010), en el cual el parámetro de regularización  $\lambda$  es estimado mediante validación cruzada en los datos de entrenamiento. Una nueva matriz  $\tilde{\mathbf{G}}$  de menor dimensión que la de la matriz original  $\mathbf{G}$  fue construida utilizando aquellas columnas de  $\mathbf{G}$  para las cuales el vector ralo solución de LASSO contenía coeficientes no-nulos. En el caso de PFBCSP, el algoritmo GSDA, que realiza selección automática de parámetros de regularización, fue implementado con criterio de parada temprana. En particular, elegimos al parámetro de parada igual a 8, es decir  $\|\hat{\beta}\|_0 \leq 8$ . Es importante notar aquí que, mientras para SFBCSP la regla de clasificación



debe ser aprendida dada la matriz de características optimizadas  $\tilde{\mathbf{G}}$ , para el caso de PFBCSP la regla de clasificación es sencillamente aprendida y aplicada utilizando el vector nuevo de características óptimas  $\mathbf{G}^* = \mathbf{G}\hat{\boldsymbol{\beta}} \in \mathbb{R}^n$ , en el cual las clases son fácilmente separables mediante una función lineal. Tal como lo sugieren en (Ang et al., 2008), la regla de clasificación fue implementada utilizando el conocido algoritmo LDA.

El análisis en múltiples ventanas temporales es una extensión directa del análisis de banco de filtros con una única ventana temporal antes descrito. En esta configuración, en lugar de utilizar un único segmento temporal de EEG pre-definido, se utilizan  $T$  segmentos de EEG de 2 s de duración con un solapamiento de 1,5 s para cada época de EEG ( $tb_1, \dots, tb_t, \dots, tb_T$ ) desde 0 hasta  $T_{\text{máx}}$  segundos luego de aplicarse la señal de MI, donde  $T_{\text{máx}}$  es la duración total de la época de MI-EEG. El análisis en múltiples ventanas junto con el banco de filtros resulta para cada época de EEG en  $q = 2KFT$  características al aplicar CSP en cada banda  $t$ - $f$  ( $q = 136$  para MI-1,  $q = 204$  para MI-2, y  $q = 170$  para MI-3). En el caso de PTFBCSP el parámetro de parada de GSDA fue fijado en  $4T$ . El desempeño de clasificación para CSP, SFBCSP, PFBCSP y PTFBCSP fue evaluado mediante validación cruzada  $10 \times 10$ . Para el análisis estadístico de resultados de clasificación, se utilizó ANOVA de un factor con bloques aleatorizados, y luego, para las comparaciones múltiples, el método de Tukey-Kramer, ambos con un nivel de significancia  $\alpha = 0,01$ .

### 6.3.2. Resultados de clasificación

Las Figuras 6.2, 6.3 y 6.4 muestran los promedios de los resultados de clasificación obtenidos por CSP, SFBCSP, PFBCSP y PTFBCSP, a lo largo de las  $10 \times 10$  validaciones para las base de datos MI-1, MI-2 y MI-3, respectivamente. Nótese que para todas las bases de datos ambos enfoques PFBCSP y PTFBCSP alcanzan la tasa de clasificación promedio más alta. En particular, el uso de múltiples segmentos temporales junto con el banco de filtros mejora notablemente la clasificación, tal como puede observarse para los sujetos 2 y 3 de la base de datos MI-1, el sujeto 4 de la base de datos MI-2 y el sujeto 1 de la base de datos MI-3. Por otro lado, existen situaciones (sujeto 3 de MI-1, sujeto 3 de MI-2 y sujeto 5 de MI-3) donde este procedimiento no resulta beneficioso.

Los resultados de clasificación promedio para cada base de datos se presentan en la Tabla 6.1. Los mejores resultados de clasificación se muestran en negrita. Se observa que ambas configuraciones superan el desempeño de los otros métodos. El análisis estadístico mostró que dichas diferencias son significativas con un  $p$ -valor  $< 0.01$ . Estas mejoras pueden

también medirse en términos de porcentajes. En el caso particular de MI-1, el método de ventana temporal única PFBCSP alcanza incrementos de clasificación de 15,2% y 3,9%, comparado con CSP y SFBCSP, respectivamente. Para el caso de la MI-2 y MI-3, estos incrementos fueron de 11,6% y 2,3%, y 10,0% y 2,9%, respectivamente. En el caso del método de múltiples ventanas temporales PTFBCSP, se encontraron incrementos de clasificación del 20,9% y 9,0%, en comparación con CSP y SFBCSP, respectivamente. Para MI-2 y MI-3 estos valores fueron 12,6% y 3,3%, y 11,4% y 4,3%, respectivamente.

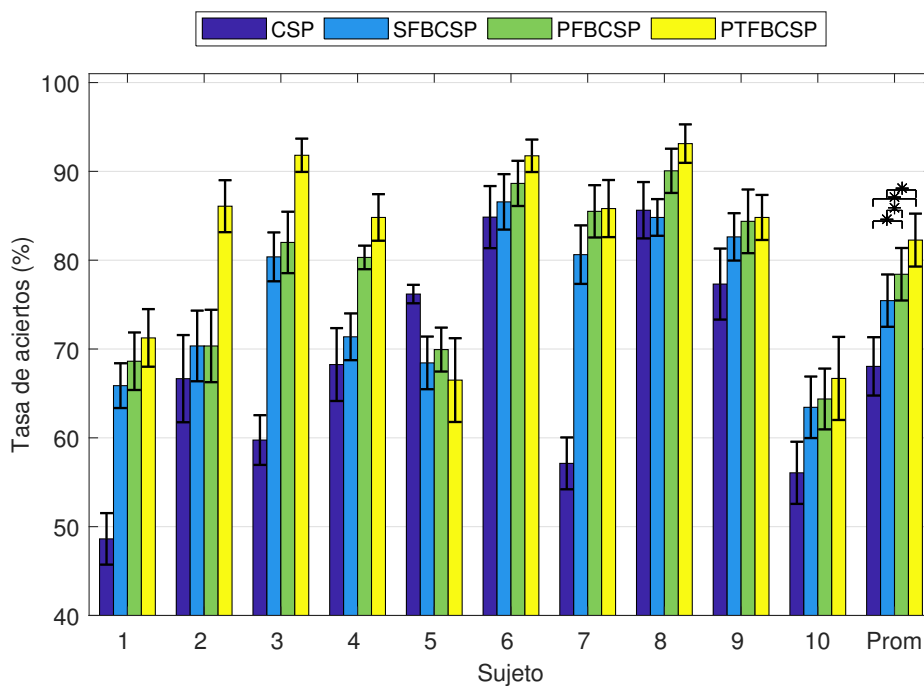


Figura 6.2: Resultados de clasificación promedio sobre las  $10 \times 10$  validaciones cruzadas obtenidas mediante CSP, SFBCSP, PFBCSP y PTFBCSP para la base de datos MI-1. Aquí “\*\*\*” denota diferencia estadística significativa entre los métodos.

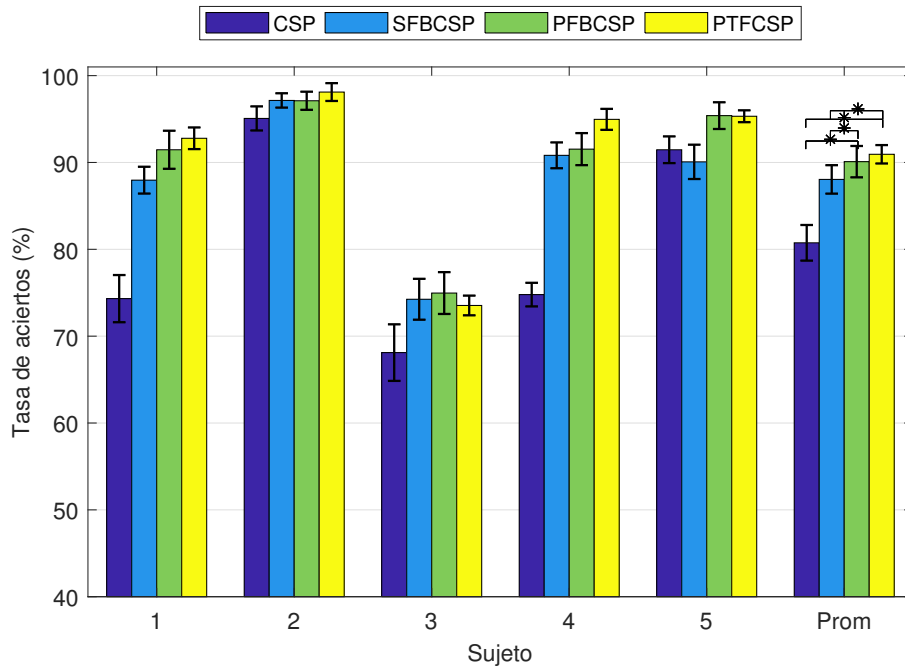


Figura 6.3: Resultados de clasificación promedio sobre las  $10 \times 10$  validaciones cruzadas obtenidas mediante CSP, SFBCSP, PFBCSP y PTFBCSP para la base de datos MI-2. Aquí “\*\*” denota diferencia estadística significativa entre los métodos.

Base de datos	CSP	SFBCSP	PFBCSP	PTFBCSP
MI-1	68,05 ( $\pm$ 3,28)	75,45 ( $\pm$ 2,95)	78,42 ( $\pm$ 2,96)	<b>82,26 (<math>\pm</math> 2,98)*</b>
MI-2	80,75 ( $\pm$ 2,05)	88,05 ( $\pm$ 1,64)	90,09 ( $\pm$ 1,81)	<b>90,94 (<math>\pm</math> 1,06)*</b>
MI-3	72,88 ( $\pm$ 3,77)	77,89 ( $\pm$ 2,66)	80,19 ( $\pm$ 2,72)	<b>81,23 (<math>\pm</math> 2,46)*</b>

Tabla 6.1: Descripción general de los valores promedio de clasificación y sus correspondientes desvíos estándar entre paréntesis. Valores expresados en porcentajes. Los mejores resultados de clasificación están resaltados en negrita. Aquí “\*\*” denota que  $PTFBCSP > SFBCSP$  con un  $p$ -valor  $< 0,01$ .

Estos resultados muestran, por un lado, que el uso de información discriminativa entre las características (PFBCSP vs. SFBCSP) mejora la clasificación (hasta en un 3,9%), y por el otro lado, que el uso de múltiples ventanas temporales (PTFBCSP vs. SFBCSP) mejora aún más la eficiencia en la detección de la intención de movimiento (mejoras del 3,3%), con mejoras muy notables cuando MI tiene que ser detectado versus condición de no-movimiento/relajación (MI-1, mejoras de un 9,0%).

Mediante la utilización del método PTFBCSP las mejoras en algunos sujetos superan el 10,0% (ver por ejemplo sujeto 2 de MI-1 en la Figura 6.2 y sujeto 1 de MI-3 en la

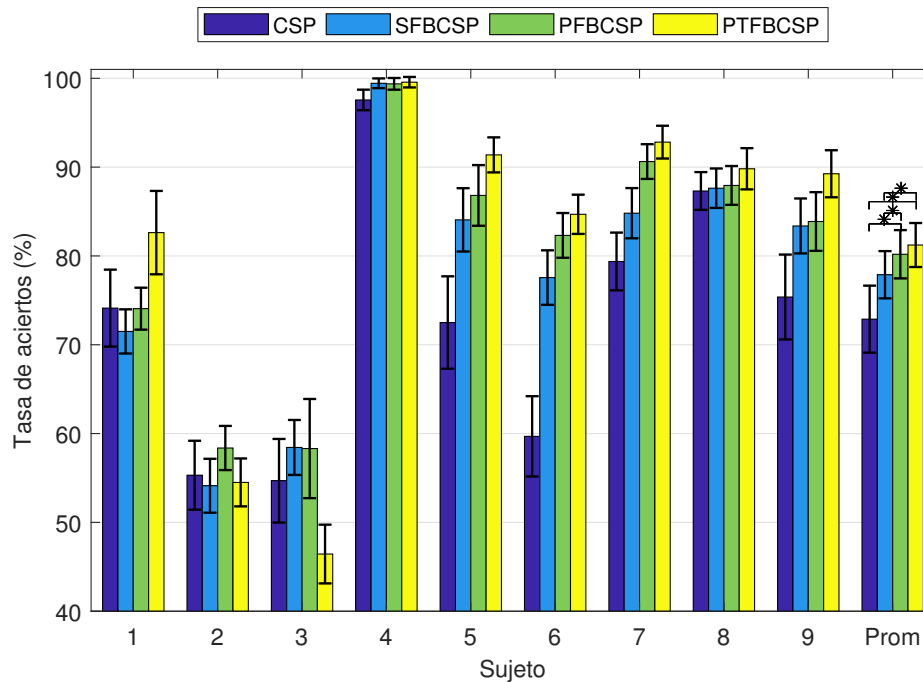


Figura 6.4: Resultados de clasificación promedio sobre las  $10 \times 10$  validaciones cruzadas obtenidas mediante CSP, SFBCSP, PFBCSP y PTFBCSP para la base de datos MI-3. Aquí “\*” denota diferencia estadística significativa entre los métodos.

Figura 6.4). Asimismo, mediante un análisis de los valores de las varianzas obtenidas por cada método en las  $10 \times 10$  validaciones cruzadas, revela que en algunos casos (MI-2), PTFBCSP resulta más robusto a los cambios en el conjunto de entrenamiento comparado con los otros métodos de ventana única considerados (ver los valores de desviación estándar en la Tabla 6.1).

### 6.3.3. Mapas topográficos de los patrones temporo-frecuenciales espaciales

El uso de CSP permite la interpretación neurofisiológica de los patrones espaciales resultantes (Haufe et al., 2014). Los patrones espaciales, es decir los vectores columna de la matriz  $\mathbf{A}$  introducida anteriormente, ilustra como las presuntas fuentes de activación se proyectan en la superficie del cuero cabelludo (Blankertz et al., 2008). El método PTFBCSP propuesto otorga un vector solución para cada sujeto con un número relativamente pequeño de coeficientes distintos de cero. Estos coeficientes no nulos representan precisamente las bandas temporo-frecuenciales seleccionadas por el método. Mediante las

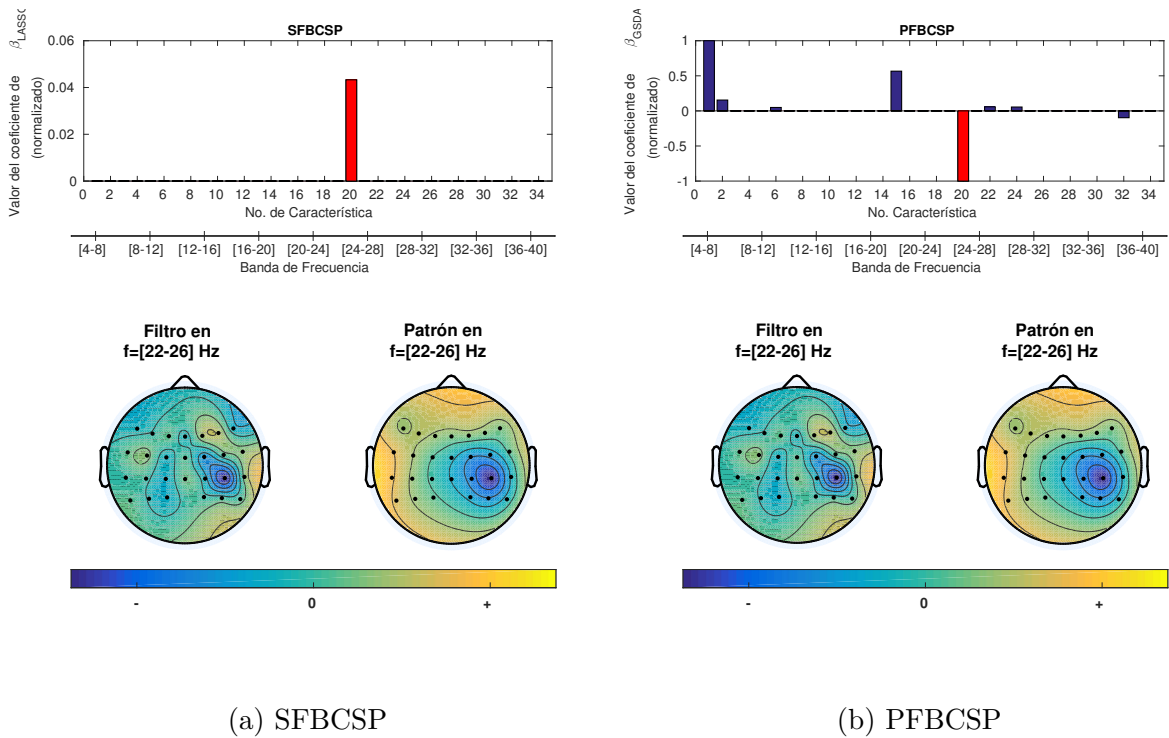


Figura 6.5: Arriba: características seleccionadas por cada método de ventana única analizado. Abajo: mapas topográficos de filtro y el patrón espacial correspondiente a la característica más relevante (resaltada en rojo) para un sujeto de la MI-1 con dominancia izquierda. a) SFBCSP, b) PFBCSP.

gráficas de los mapas topográficos tanto del filtro como del patrón espacial asociado a la característica más relevante, la correlación neurofisiológica de la solución hallada con la señal de interés puede ser analizada. Las Figuras 6.5 y 6.6 muestran, para los métodos de ventana temporal única (SFBCSP y PFBCSP) y para el método de múltiples ventanas temporales PTFBCSP, respectivamente, el correspondiente vector solución junto con el mapa topográfico del filtro y del patrón espacial asociado a la característica más relevante. Estas gráficas corresponden a una época aleatoria del sujeto 2 (dominancia izquierda) de la base de datos MI-1.

En la Figuras 6.5 y 6.6 se ve, por un lado, que para los métodos de ventana única la información frecuencial se encuentra localizada entre 22 y 26 Hz, mientras que para el método de múltiples ventanas temporales la banda frecuencial más relevante varía según sea el segmento temporal. Es notorio como ambos métodos de ventana única seleccionan la misma característica como la más relevante, la cuál (mediante inspección visual de su correspondiente patrón espacial) evidencia la correlación neuronal con la imaginación del gesto de abrir y cerrar la mano (activación contralateral al movimiento imaginado). Por

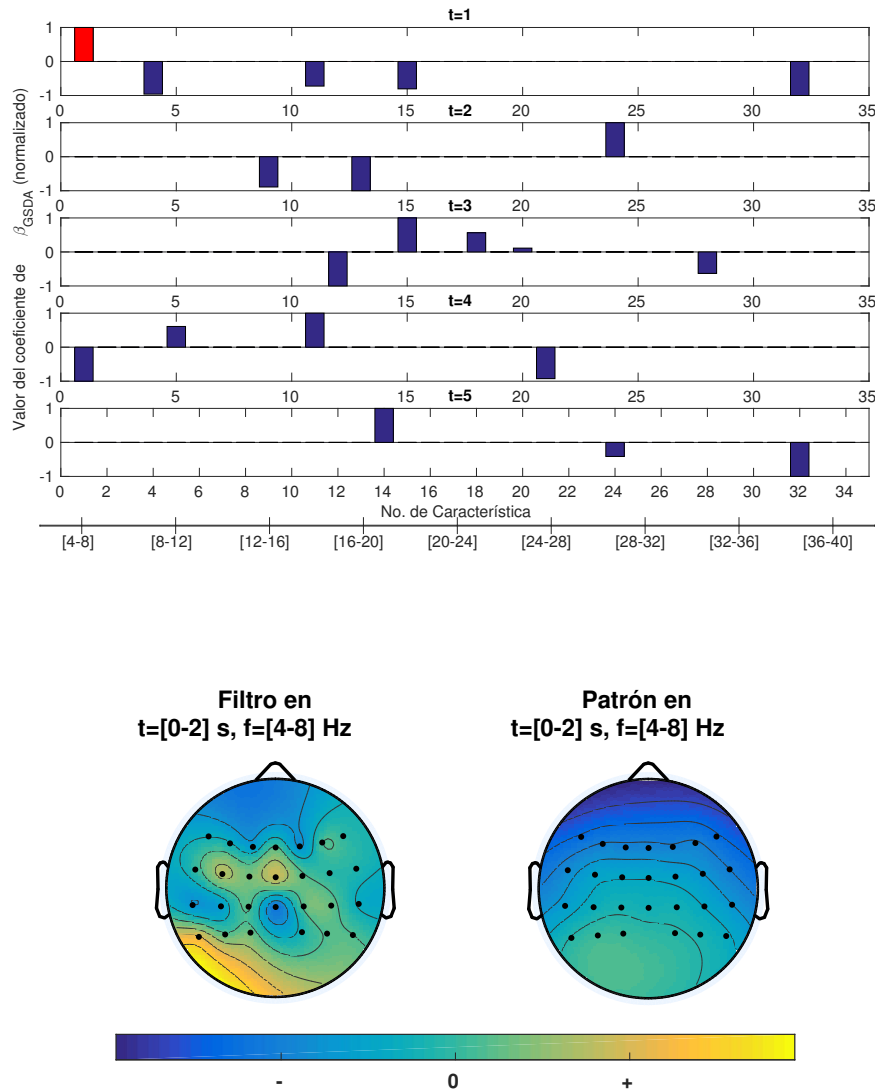


Figura 6.6: Arriba: características espectro-temporales espaciales seleccionadas por el método propuesto de banco de filtros y múltiples ventanas temporales en cada segmento temporal  $t = 1, 2, \dots, 5$  (uno por fila). Abajo: mapa topográfico del filtro y del patrón espacial asociado a la característica más relevante (resaltada en rojo) para un sujeto de la MI-1 con dominancia izquierda.

otro lado, los mapas topográficos para el método de múltiples ventanas temporales parecen representar la condición de relajación. Es importante mencionar aquí, que desde el punto

de vista cognitivo, la condición de relajación/no-movimiento compromete una variedad de estados mentales que el usuario de la BCI podría estar realizando en contraposición a la MI. De hecho, suponemos que éste es el motivo por cual todos los métodos evaluados resultan en tasas de clasificación más bajas para la MI-1 en comparación a las otras dos bases de datos consideradas.

## 6.4. Discusiones

En este capítulo hemos propuesto un método de análisis de señales de EEG para seleccionar de manera óptima y automática tanto las bandas frecuenciales como los segmentos temporales para mejorar la detección de la intención de movimiento. Este método consiste en descomponer la señal de EEG multicanal en  $T$  segmentos temporales y  $F$  bandas frecuenciales. Mediante el uso combinado de CSP y GSDA, el primer filtro espacial asociado a cada clase es utilizado para la extracción de características, y luego la selección de características y clasificación es realizada en simultáneo con selección automática de parámetros. Mediante este método, hemos sido capaces de validar la potencialidad de GSDA en otro contexto dentro de BCI. Los resultados muestran que mediante la optimización sujeto-específica tanto de las bandas temporales como de las frecuenciales, pueden obtenerse mejoras de clasificación de hasta un 20% en comparación con métodos actuales del estado del arte para detección de MI.

De las dos diferentes configuraciones del método PTFBCSP propuesto, los resultados claramente indican que el uso de múltiples ventanas temporales mejora la detección de MI, más aún cuando las tareas de MI deben ser detectadas contra tareas no específicas, como lo es la condición de relajación. Este hecho muestra la capacidad de generalización del método, el que, de un número muy grande de posibles características, selecciona aquel subconjunto óptimo desde el punto de vista de la discriminabilidad. Como desventaja, a pesar de que la ventana temporal óptima para cada sujeto puede ser seleccionada como aquella con la mayor cantidad de coeficientes distintos de cero, (por ejemplo, segmento 1 en la Figura 6.6), se requiere de la longitud total de la época de MI para analizar la señal. Esto supone una limitación para aplicaciones de tiempo real que requieran de respuestas rápidas. En el caso de que se desee un tiempo de decisión más corto, la variante de ventana única PFBCSP representa un buen compromiso entre la longitud temporal de cada época de EEG luego del comienzo de MI y los valores de clasificación.

Los resultados de clasificación aquí encontrados para las bases de datos de acceso público (MI-2 y MI-3) pueden compararse con los resultados hallados por dos métodos de

análisis temporo-frecuenciales del estado del arte. En particular, en (Ang et al., 2012b) los resultados de clasificación sobre  $10 \times 10$  validaciones cruzadas con la MI-3 fueron reportados en  $78,25 \pm 1,12\%$ , mientras que para nuestro método PTFBCSP los valores análogos fueron  $81,23 \pm 2,46\%$ . Una comparación similar se puede realizar con el método propuesto en (Miao et al., 2017), en el cual, luego de una apropiada selección sujeto-dependiente de canales, los resultados de clasificación reportados para la MI-2 fueron  $92,66 \pm 4,44\%$ . Si bien este último supone una pequeña mejora en los valores de clasificación (que puede deberse a la selección de canales realizada) comparada con PTFBCSP ( $90,94 \pm 1,06\%$ ), el desvío estándar reportado es mucho más grande, volviendo a PTFBCSP un método más robusto y estable.

Si bien hemos encontrados muy buenos resultados de clasificación ( $> 80\%$ ) junto con varianzas bajas a lo largo de las 10 repeticiones de la validación cruzada de 10 particiones, para todas las bases de datos utilizadas, es necesario estudiar el comportamiento del método propuesto en escenarios de tiempo real, en los que los tiempos de decisión deben ser lo más breve posibles. Asimismo, estos resultados pueden ser mejorados si se estudian nuevas y diferentes configuraciones de los electrodos, así como métodos de selección de los parámetros de regularización sujeto-dependientes. De todas maneras, consideramos que los resultados aquí expuestos constituyen un importante avance hacia la detección de la intención de movimiento para la rehabilitación de pacientes con severa pérdida de control motor de uno de sus miembros superiores.





## Capítulo 7

# Hacia la aplicación final: evaluación en tiempo real de una BCI basada en MI

“En todas las artes, como en todas las artesanías, el talento sirve de bien poco si no va acompañado por la experiencia, que se logra a fuerza de modestia y aplicación.”

---

Patrick Süskind, El perfume

En este capítulo describiremos brevemente el trabajo realizado para la toma de datos y evaluación en tiempo real del método desarrollado para la detección de intención de movimiento. Motivados en la aplicación de BCI basadas en MI para la rehabilitación de pacientes con severa pérdida del control voluntario de alguna de sus extremidades superiores, diseñamos, desarrollamos y evaluamos dicho método en tiempo real tanto con sujetos sanos como con un paciente un accidente cerebro vascular.

### 7.1. Introducción

Según la Organización Mundial de la Salud, 15 millones de personas en todo el mundo sufren accidentes cerebro vasculares (ACV) cada año, provocando diversos grados de paresia en las extremidades superiores en aproximadamente un tercio de ellas ([Benjamin et al., 2017](#)). En Argentina, el Ministerio de Salud de la Nación indica que esta patología afecta anualmente a más de 80.000 habitantes. Los procedimientos tradicionales de rehabilitación están fuertemente orientados a que el paciente use más a menudo su mano

afectada, ya sea a través de movimientos activos (controlados por el sujeto), o pasivos (por ejemplo, realizados mediante la ayuda de un robot) (Hesse et al., 2005; Wolf et al., 2006). Para aquellos pacientes con nula o muy poca actividad muscular remanente en la mano afectada, el uso de BCI basadas en MI constituye un abordaje novedoso e integrador para la rehabilitación neuromotora de estos pacientes (Millán et al., 2010). Como se describió anteriormente (ver Sección 2.2.2), mediante una MI-BCI la actividad cerebral asociada al deseo/imaginación de un cierto movimiento es directamente transformada en señales de control de dispositivos externos, tal como una computadora o un brazo robótico (Frolov et al., 2017). Terapias de rehabilitación en las que el paciente, mediante una MI-BCI controla una ortesis mecánica (exoesqueleto) colocada en la mano paralizada, cierra el circuito transformando la intención en ejecución del movimiento. Se ha demostrado que la retroalimentación visual y sensorial que percibe un sujeto puede mejorar las funciones motoras de la mano, y también, probablemente, la plasticidad cerebral en la zona dañada (Broetz et al., 2010; Subramanian et al., 2010).

Para que las terapias de rehabilitación sean más accesibles a sus potenciales usuarios, realizables en el hogar y/o enmarcadas en un tratamiento fisioterapéutico complementario, es entonces necesario disponer, por un lado, de equipos de adquisición transportables, confiables y comparables a equipos de grado clínico, y por otro lado de algoritmos de traducción eficientes y versátiles, capaces de ajustarse de manera rápida y automática a las posibles variaciones en las mediciones de la actividad cerebral.

En el capítulo anterior propusimos un método para mejorar la detección de la intención de movimiento. Los resultados hallados mediante los experimentos de validación cruzada indican que el método de ventana única PFBCSP es un buen candidato para la detección en tiempo real de MI. En este capítulo describiremos el trabajo realizado para evaluar la factibilidad y aplicabilidad de dicho método en tiempo real para una BCI basada en MI con aplicaciones en rehabilitación. En una primera etapa dicha evaluación fue realizada con sujetos sanos. Luego, dado que los patrones de activación cerebral pueden ser diferentes para pacientes con alteraciones neurológicas, todo el diseño de la MI-BCI fue transferido y evaluado en los posibles usuarios finales de dicho sistema: pacientes de ACV.

## 7.2. Descripción del experimento

El objetivo principal de las MI-BCIs en el contexto de rehabilitación motora funcional es activar las conexiones cerebrales sensoriomotoras. Por ello, es crucial que los usuarios de las BCIs realicen una tarea mental desde una perspectiva propioceptiva focalizada en

las sensaciones asociadas a la realización de un cierto movimiento (MI cinestético). Más aún, para lograr una mejor simulación mental de un movimiento, la tarea de MI debe ser tomada de una actividad simple de la vida diaria, tal como lo es el gesto de abrir y cerrar la mano (Millán et al., 2010).

Teniendo en cuenta dichas características se diseñó un protocolo de experimentación de MI cinestético vs. relajación. Para posteriormente analizar la variabilidad intra-sujeto dada por los cambios en las condiciones pre-fisiológicas, se realizó la toma de datos en dos sesiones (días) diferentes con una máxima separación entre sesión y sesión de 5 días. Cada sesión contemplaba cuatro rondas o corridas. Cada corrida de experimentación constaba de 20 repeticiones de imaginación motora cinestética intercaladas aleatoriamente con 20 épocas de no imaginación/relajación. Las primeras dos rondas eran utilizadas para entrenar el método PFBCSP propuesto para la detección de la intención de movimiento. Si bien, ningún tipo de realimentación o devolución era realizada para con el sujeto, las segundas dos rondas fueron utilizadas para la evaluación en tiempo real de la detección de MI.

Para obtener información adicional de los participantes, se diseñaron y utilizaron dos cuestionarios: uno de atención y otro de imaginería motora (*Kinesthetic and Visual Imagery Questionnaire*, KVIQ (Malouin et al., 2007)). Para la elección de los posibles sujetos de experimentación, se consideraron criterios de inclusión/exclusión detallados en la Tabla 7.1

Para el reclutamiento de sujetos sanos se realizó un llamado de voluntarios por medio de correo electrónico y comunicación verbal a estudiantes pertenecientes al “Balgrist Campus”<sup>1</sup>, Zurich, Suiza. Para el caso de los pacientes, esta investigación fue realizada en conjunto con el Instituto de Rehabilitación “Center for Neurological Rehabilitation”<sup>2</sup> (Cereno, Vitznau, Suiza).

El protocolo de experimentación aquí descrito fue evaluado y aprobado por la Comisión de Ética de Zurich, Suiza, BASEC-Nr. Req-2017-00631.

### 7.2.1. Protocolo de estimulación

El protocolo de estimulación diseñado, que se encuentra esquemáticamente representado en la Figura 7.1, está basado en el método estándar de estimulación propuesto por el prestigioso grupo de investigación de Graz, Austria (Pfurtscheller y Neuper, 2001). Cada época del experimento comenzaba con una cruz verde en el medio de la pantalla

<sup>1</sup><http://www.balgristcampus.ch/en/home/>

<sup>2</sup><https://www.cereneo.ch/en/home.html>

SUJETOS SANOS	
Criterios de inclusión	Criterios de exclusión
<ul style="list-style-type: none"> <li>✓ Ser mayor de edad.</li> <li>✓ Tener la habilidad de dar consentimiento informado y entender comandos simples.</li> <li>✓ No poseer disfunciones motoras.</li> </ul>	<ul style="list-style-type: none"> <li>× Haber abusado de drogas y/o alcohol.</li> <li>× Poseer una enfermedad clínica significativa del CNS o del aparato motor.</li> </ul>
PACIENTES	
Criterios de inclusión	Criterios de exclusión
<ul style="list-style-type: none"> <li>✓ Ser mayor de edad.</li> <li>✓ Tener la habilidad de dar consentimiento informado y entender comandos simples.</li> <li>✓ Poseer discapacidades severas de miembro superior, reflejado en valores clínicos estándar evaluadas por un equipo terapéutico.</li> </ul>	<ul style="list-style-type: none"> <li>× Haber abusado de drogas y/o alcohol.</li> <li>× Poseer lesiones en la corteza motora primaria, identificada mediante documentación médica.</li> <li>× Pacientes que sufran de manera recurrente de ACV y/o hayan sido sometidos a cirugía cerebral.</li> <li>× Pacientes con ataxia, distonía, temblor en miembro superior o mioelopatía cervical previa.</li> </ul>

Tabla 7.1: Criterios de inclusión y exclusión para sujetos sanos (arriba) y pacientes de ACV (abajo).

( $t = -3$  s). Para llamar la atención del sujeto/paciente, una señal audible era presentada dos segundos más tarde ( $t = -1$  s). Al comienzo de una época ( $t = 0$  s), si se presentaba una flecha roja en la pantalla, el sujeto/paciente debía imaginar/activamente realizar el movimiento de abrir y cerrar su mano dominante/afectada por un período de tiempo de 4 s en un ciclo de apertura-cierre de 1 Hz (clase MI). Por el contrario, si no aparecía la flecha roja (sólo cruz verde en la pantalla), la persona debería simplemente relajarse y no imaginar movimiento alguno (clase relajación) por el mismo período de tiempo. Para evitar el acostumbramiento y/o agotamiento por parte del sujeto, al finalizar cada repetición, y hasta el comienzo de la siguiente, la pantalla se ponía negra por un período de tiempo aleatorio (entre 2,5 y 4,5 s). Asimismo, un período de tiempo más largo ( $> 2$  min) luego de cada ronda permitía al sujeto descansar.

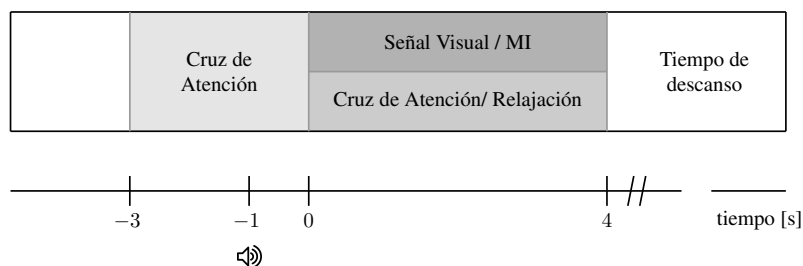


Figura 7.1: Representación esquemática de una época del experimento con la referencia de los tiempos del protocolo, en segundos.

### 7.2.2. Software y hardware utilizados

Para la toma de datos se utilizó un amplificador eego<sup>TM</sup>rt Ant Neuro amplifier (Ant Neuro, Holanda). Los datos de EEG fueron adquiridos utilizando 64 electrodos no-invasivos localizados en concordancia con el sistema internacional 10-20, utilizando los electrodos CPz y AFz como referencia y tierra, respectivamente (ver Figura 7.2). Las señales fueron muestreadas a 512 Hz y filtradas entre 0,5 Hz y 40 Hz con un filtro Butterworth de orden 3 pasa-banda mediante el software de adquisición. Para monitorear que el sujeto no realice movimiento físico/real de su mano, electromiografía (EMG) de superficie de dos músculos de la mano y de dos músculos del antebrazo fue adquirida en simultáneo con el EEG (Noraxon TeleMyo DTS, Noraxon, Estados Unidos).

El protocolo de estimulación junto con el registro de las señales de EEG se realizó mediante el OpenVIBE<sup>3</sup>. OpenVIBE es un software libre multiplataforma que permite el

<sup>3</sup><http://openvibe.inria.fr/>

diseño, evaluación y uso de BCIs. Mediante dicha plataforma se pueden registrar señales biomédicas, realizar técnicas de procesamiento básico de señales y visualizar las mismas en tiempo real. Tiene la ventaja de ser compatible con diferentes sistemas de adquisición de señales.

Para registrar y visualizar las señales de EMG se utilizó el software Noraxon Myo-Research<sup>4</sup>. Las señales de EMG fueron registradas de manera sincronizada con el comienzo de cada ronda de EEG. Para ello se utilizó la librería *Lab Streaming Layer*<sup>5</sup> (LSL) junto con el software Matlab<sup>®</sup>. Esta librería fue utilizada para visualizar en Matlab las señales y los marcadores enviados por y desde OpenVIBE (comienzo de corrida, tipo de tarea MI, fin del experimento, etc.). En las rondas de evaluación, en las que el clasificador debía detectar en tiempo real la presencia de MI, también gracias al uso de dicha librería, el registro y clasificación de los segmentos de EEG fue realizado directamente en Matlab.

En la Figura 7.3 se representa esquemáticamente el montaje del experimento con software y hardware utilizados.

### 7.2.3. Adquisición de datos

Considerando los criterios de inclusión y exclusión, y en función de los pacientes disponibles en la clínica, se realizaron registros de once (11) sujetos sanos (3 mujeres, 4 de dominancia izquierda, edad promedio  $25,45 \pm 2,50$  años) y de un (1) paciente sobreviviente a un ACV (mujer, 42 años, mano derecha afectada).

Para el caso de los sujetos sanos, cada sujeto se encontraba cómodamente sentado frente a un monitor de computadora con ambos brazos descansando sobre un escritorio. Para evitar que el sujeto mire su mano e interfiera en el ejercicio cinestético mental de imaginar el movimiento, su brazo dominante con los electrodos de EMG sujetos a él, se

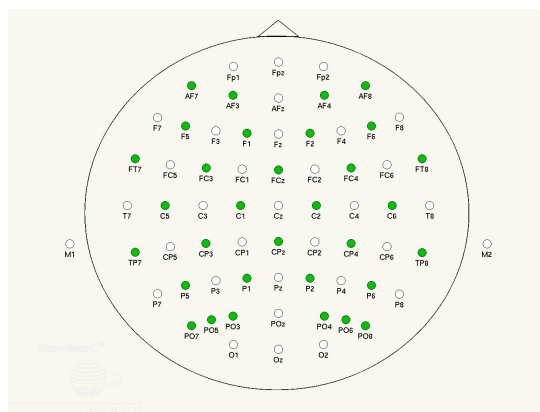


Figura 7.2: Localización de los 64 electrodos utilizados para la toma de datos. Nótese que CPz ha sido utilizado como electrodo de referencia y AFz como tierra.

<sup>4</sup><https://www.noraxon.com/our-products/myoresearch/>

<sup>5</sup><https://github.com/sccn/labstreaminglayer>

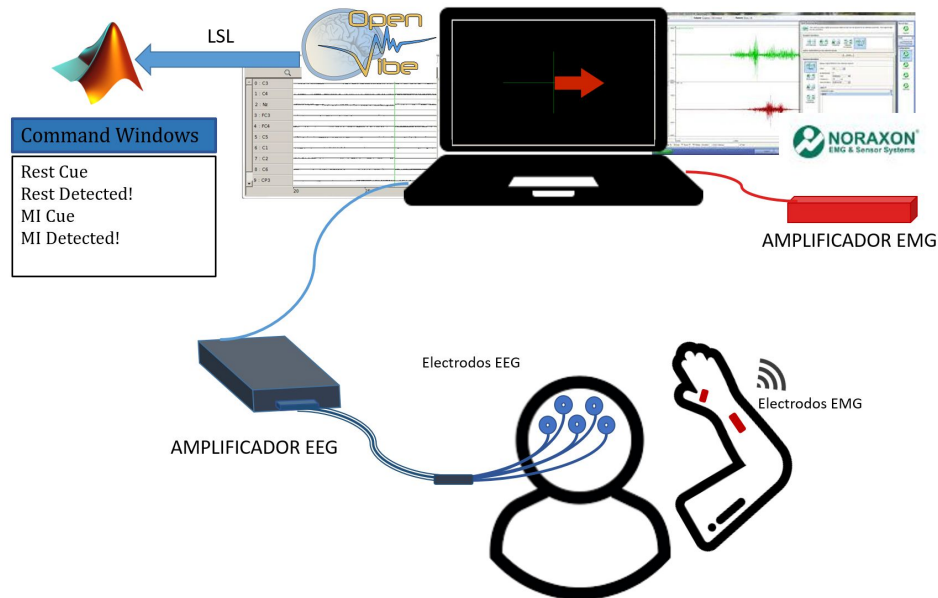


Figura 7.3: Descripción esquemática del experimento. Electrodo, equipo de adquisición y softwares utilizados para el registro de cada señal biomédica considerada.

ubicó dentro de una caja de cartón (ver Figura 7.4). La máxima contracción voluntaria (MVC, por sus siglas en inglés), es decir la máxima fuerza que una persona puede realizar en un ejercicio isométrico específico, fue adquirida al comienzo de cada sesión para tener luego una referencia en el análisis de la actividad muscular dentro de las épocas de EEG. En el caso del paciente, dado que el experimento fue realizado en la misma clínica y como parte de su sesión de rehabilitación, la actividad muscular no pudo ser registrada. Por otro lado, a este sujeto, en lugar de pedirle que imagine la realización del movimiento, se le pidió que efectivamente realice la intención de movimiento de su mano afectada.

Las rondas de calibración (primeras dos rondas de cada sesión), contenían señales de EEG con su correspondiente etiqueta de clase (dada por el protocolo de estimulación). Este conjunto de datos (80 épocas de EEG, 40 MI/40 relajación), se utilizó para entrenar el método de detección PFBCSP implementado en Matlab<sup>®</sup>. Las segundas dos rondas del experimento fueron idénticas para el sujeto, excepto que, en Matlab nuestro clasificador predecía en tiempo real cada época de EEG extraída. El sujeto no recibió retroalimentación alguna de su desempeño por parte del sistema, para no influenciar el experimento y tener la posibilidad de poder luego evaluar otros enfoques o mejoras del método. La segun-



da sesión fue idéntica a la primera, lo que permitirá evaluar el desempeño del clasificador entrenado y la robustez ante las variaciones intra-sujeto.



Figura 7.4: Montaje del experimento para un sujeto con dominancia izquierda. Se muestran los sistemas de adquisición de EEG y EMG, como así también el protocolo de estimulación con la señal visual para una época de MI

## 7.3. Resultados

### 7.3.1. Análisis de la señal de EMG

La adquisición de señales de EMG sólo pudo ser realizada para los sujetos sanos. Por lo cual, el siguiente análisis sólo es válido para los datos de EEG provenientes de los 11 sujetos sanos.

Las señales de EMG fueron analizadas para poder detectar (si existiesen), para cada sujeto, cuáles y cuántas épocas de EEG estaban contaminadas por ruido muscular. Además de la inspección visual de las señales de EMG, una manera estándar de detectar la activación muscular consiste en determinar el momento en que la señal de EMG

supera un cierto umbral (Hodges y Bui, 1996). La envolvente de la señal de EMG puede estimarse aplicando a la señal de EMG centrada (media cero) y rectificada un filtro Butterworth pasa-bajos de orden 4 sin corrimiento de fase con frecuencia de corte en 50 Hz. Los segmentos de EMG se extrajeron en correspondencia con el comienzo de una época de EEG. Se utilizaron ventanas móviles de 500 ms a lo largo de dichos segmentos para analizar la actividad muscular. Para cada época y para cada uno de los cuatro músculos considerados, se etiquetó como época contaminada a aquella realización cuyo valor medio en la ventana móvil de EMG superaba al 20 % del correspondiente valor máximo de la envolvente del MVC. Para uno de los sujetos, más del 50 % de las épocas en una ronda estaban contaminadas por actividad muscular, razón por la cual los resultados de clasificación en tiempo real para este sujeto no serán reportados.

### 7.3.2. Resultados de clasificación en tiempo real

Con las señales de EEG correspondientes a las dos primeras corridas, para cada sujeto y para cada sesión, se entrenó el algoritmo de detección de intención de movimiento basado en el método PFBCSP. Para ello, y siguiendo la misma metodología explicada en la Subsección 6.3, se utilizaron 28 electrodos cubriendo el área sensoriomotora y segmentos de EEG extraídos entre 0,5 y 2,5 s luego de la presentación del estímulo visual ( $t = 0$  s). Para cada segmento de EEG se aplicó un banco de filtros de 17 sub-bandas frecuenciales entre 4-40 Hz con ancho de banda 4 Hz y solapamiento 2 Hz. Los dos primeros filtros aprendidos mediante CSP en cada sub-banda fueron utilizados para extraer 2 características por banda frecuencial. Mediante GSDA se realizó la selección y clasificación de características. Al finalizar el proceso de entrenamiento, para cada sujeto y para cada sesión, se contaba con un “modelo”, constituido por las matrices CSP aprendidas para cada sub-banda, la media y desvío muestral (necesarios para normalizar los datos antes de utilizar GSDA), el vector solución  $\beta$  de GSDA y los parámetros del hiperplano de separación correspondiente a la regla de clasificación lineal. Estos parámetros aprendidos, permiten entonces la clasificación en tiempo real de los segmentos de EEG de manera rápida ( $< 40$  ms) y sujeto-específica.

La función de distribución acumulativa binomial (BCDF, por sus siglas en inglés), provee un mínimo umbral en los resultados de clasificación, necesario para el control eficiente de una BCI (Combrisson y Jerbi, 2015; Müller-Putz et al., 2008). Para un problema de clasificación balanceado, el nivel teórico mínimo de clasificación es 50 %. Este valor solo se alcanza para un número infinito de realizaciones. En la práctica, dicho nivel mínimo

depende de la cantidad de observaciones disponibles en el conjunto de datos de entrenamiento. Mediante el uso de BCDF con un nivel de confianza del 95% y basado en el número de épocas de EEG por sujeto (80 para dos primeras rondas), el mínimo umbral de clasificación encontrado fue de 58,75%.

Las Figuras 7.5 y 7.6 muestran los resultados de clasificación en cada ronda de evaluación para ambas sesiones obtenidas por los 10 sujetos sanos considerados y por el paciente con ACV, respectivamente. La línea roja denota el mínimo umbral de clasificación calculado mediante BCDF. Nótese que el valor de clasificación promedio en las dos rondas de evaluación para los sujetos sanos fue de  $72,31 \pm 12,32$ , mientras que para el paciente este valor fue de  $86,25 \pm 8,84$ .

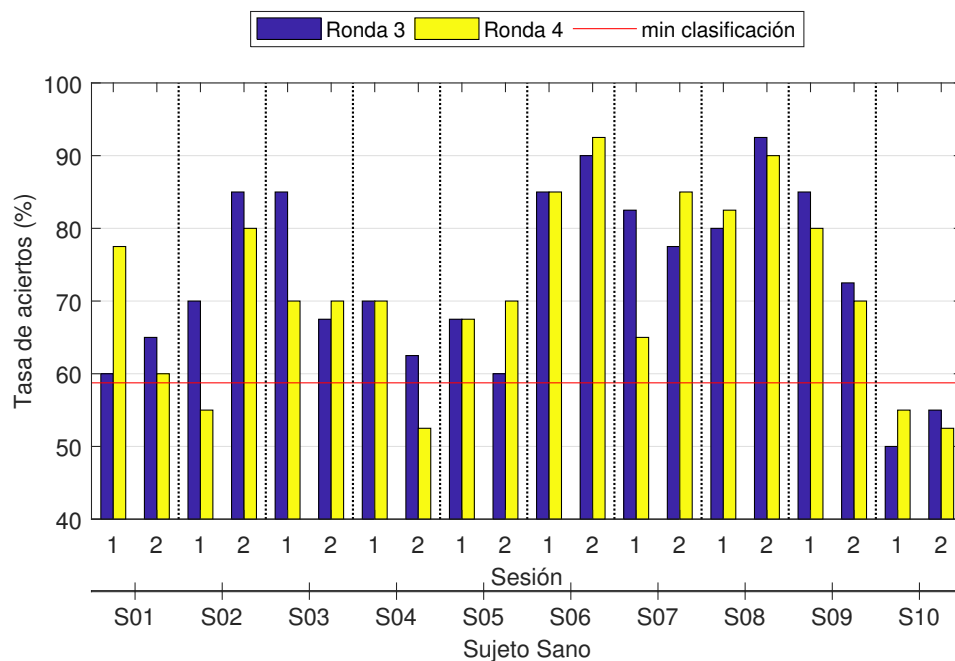


Figura 7.5: Resultados de clasificación en cada ronda de evaluación para los 10 sujetos sanos. Se muestran las dos sesiones por sujeto. Las líneas discontinuas demarcan la separación entre sujeto y sujeto. La línea roja denota el mínimo valor de clasificación para controlar una BCI.

Como se puede observar en ambas figuras, la variabilidad intra-sujeto (de ronda a ronda y de sesión a sesión) es notable, al punto tal que para algunos sujetos la clasificación de algunas de sus rondas de evaluación se encuentra por debajo del mínimo umbral, cercanas a 50%. En el caso particular del sujeto S10, este no alcanza dichos niveles mínimos de clasificación para controlar de manera confiable una BCI. Finalmente, obsérvese que la

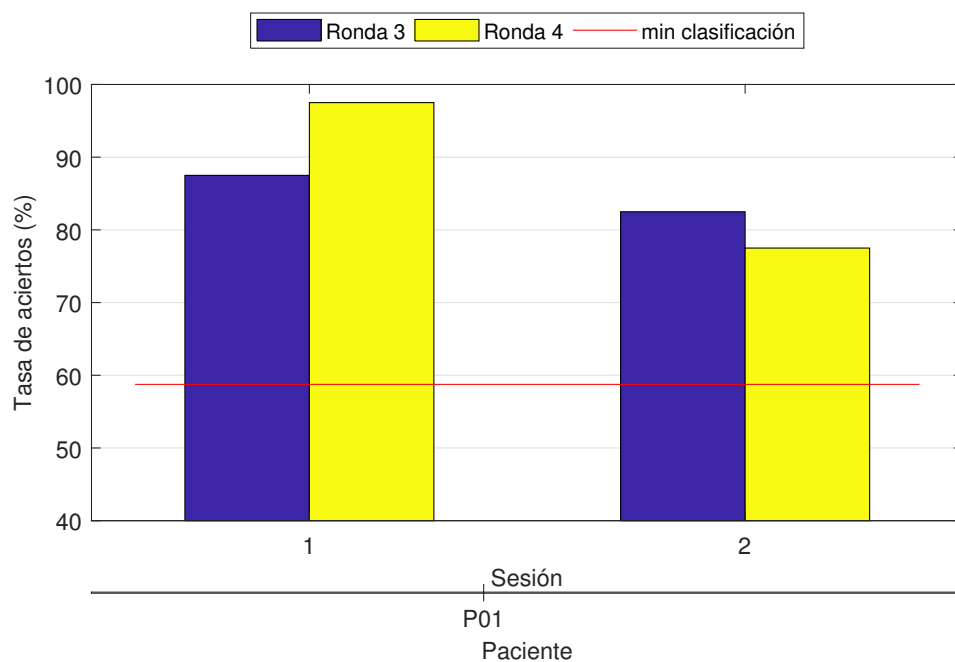


Figura 7.6: Resultados de clasificación en cada ronda de evaluación para el paciente con ACV. La línea roja denota el mínimo valor de clasificación para controlar una BCI.

evaluación del método en el paciente con ACV es altamente prometedora, alcanzando niveles de clasificación de hasta un 97 %.

## 7.4. Discusiones

En este capítulo hemos mostrado el desarrollo y posterior evaluación de la implementación en tiempo real de un método para la detección de la intención de movimiento. Para ello fue necesario diseñar un protocolo de experimentación con aplicaciones en terapias a rehabilitación. Asimismo, para el registro de las señales (EEG y EMG) se han elegido equipos de adquisición confiables, transportables, reconocidos y aceptados por la comunidad científica. En esta instancia, a modo de utilizar las funciones ya implementadas en Matlab, fue necesario contar con tecnologías que permiten utilizar dicha herramienta en tiempo real, estableciendo comunicaciones estables entre los diferentes paquetes de software y sistemas utilizados. Mediante LSL no sólo fuimos capaces de “leer” directamente desde Matlab las señales de EEG, para inmediatamente clasificarlas en tiempo real, sino que además se pudieron sincronizar de manera automática los registros de EMG con el

comienzo de la toma de datos de EEG.

El uso del mínimo umbral de clasificación permite evaluar qué sujeto es capaz de controlar una BCI. Existe evidencia de que entre un 15 y un 30 % de la población no es capaz de controlar una BCI basada en SMR, aún luego de largas y repetidas sesiones de entrenamiento (Dickhaus et al., 2009). Por otro lado, para personas sin conocimiento previo en MI-BCIs son necesarias entre 7 y 8 sesiones de entrenamiento para poder modular la actividad cerebral de manera confiable (Zhang et al., 2015a). Esto evidencia que el entrenamiento tanto del sujeto como del sistema BCI son necesarios de manera continua y adaptiva. Sumado a esto, para que un sujeto sea capaz de aprender a modular su actividad cerebral, la retroalimentación recibida por parte del sistema debe ser confiable para evitar la frustración del usuario. Por este motivo, es que en estos experimentos los sujetos no han recibido retroalimentación alguna de su desempeño en tiempo real, y sólo al finalizar la sesión, si ellos así lo deseaban, se le indicaba el porcentaje de clasificación alcanzado en cada ronda de evaluación.

Los resultados aquí expuestos permiten analizar, en una primera instancia, la factibilidad de la utilización del método propuesto para detección de MI. Son resultados preliminares, obtenidos mediante una sencilla propuesta de entrenamiento y validación fija. Estos resultados son muy prometedores, aunque podrían mejorarse aún más al utilizar técnicas de adaptabilidad del método entrenado (Shenoy et al., 2006) y enfoques de transporte óptimo (Courty et al., 2014), por nombrar algunos. Asimismo es necesario evaluar el número óptimo de observaciones de entrenamiento para cada sujeto y la posibilidad de aplicar técnicas de transferencia de modelos de sesión a sesión (Ang et al., 2012a).

La evaluación en la clínica del enfoque propuesto es sumamente optimista. Si bien, dada la lesión cerebral de los pacientes con ACV, la señal a detectar implica la intención efectiva de la realización del movimiento y no la simulación mental del mismo, los resultados de clasificación están siempre cercanos o son superiores al 80 %. Esto significa un gran avance hacia la utilización de terapias de rehabilitación para recuperación motora funcional basadas en MI-BCIs.

## Capítulo 8

# Conclusiones y trabajos futuros

“La diferencia esencial entre la emoción y la razón es que la emoción te lleva a la acción y la razón a las conclusiones.”

---

Donald Calne

Las interfaces cerebro-computadora permiten una vía de comunicación nueva y alternativa entre el cerebro de una persona y el mundo exterior. Independientemente del paradigma utilizado para establecer dicha comunicación, las BCIs dependen fuertemente del algoritmo utilizado para transformar las entradas (señales de EEG) en salidas de control. En esta tesis, luego de detectar los principales desafíos que actualmente deben enfrentar las BCIs, se diseñaron métodos discriminativos basados en técnicas de regularización con penalización mixta. El método GSDA, concebido como un método integrado en el que la selección de características es simultáneamente realizada junto con la clasificación, fue evaluado como herramienta discriminativa tanto para BCIs basadas en ERPs, como para BCIs basadas en MI. En estas últimas, GSDA estaba precedido por un necesario análisis de la señal en banco de filtros y/o segmentos temporales.

Los métodos propuestos y su eficiencia en la decodificación de la actividad cerebral fueron principalmente evaluados en experimentos de validación cruzada y comparados con los algoritmos competitivos del estado del arte. Como resultado de dicho análisis observamos que la introducción de la información discriminativa en GSDA impacta positivamente en la rareza de la solución, en la velocidad de convergencia del método, y finalmente en la clasificación de las señales, comparado al caso de no utilización de dicha información (GSDA vs. SDA). Los experimentos en escenarios adversos de pequeños datos de entrenamiento, mostraron la potencialidad de utilizar dicha técnica para disminuir los tiempos de calibración de BCIs basadas en ERPs. Por otro lado, mediante el método PTFBCSP,

basado en bandas temporo-frecuenciales y GSDA, extendimos el uso del clasificador propuesto a la resolución de otro problema binario dentro de las BCIs, mostrando también aquí la eficiencia de dicho método.

El método GSDA fue siempre implementado mediante algún criterio de parada temprana, utilizando un parámetro como cota superior de la norma- $\ell_1$  o de la norma- $\ell_0$  de la solución. Dicha elección estuvo siempre basada en conocimiento previo del problema, y se imponía de manera no específica para todos los casos de estudio (sujetos de cada base de datos). Una manera de mejorar aún más los resultados plasmados en esta tesis es mediante la selección automática y sujeto-específica de dichos parámetros.

Además de este análisis comparativo, fuimos capaces de diseñar, implementar y evaluar una MI-BCI en tiempo real. Dicha BCI estaba fuertemente motivada en aplicaciones para rehabilitación de pacientes de ACV con disfunciones motoras severas. Además del diseño del protocolo (MI vs. relajación), tanto para los experimentos fuera de línea como para la implementación en tiempo real, se utilizaron 28 electrodos cubriendo la corteza sensoriomotora de los 64 electrodos disponibles. Es importante mencionar aquí que mejores resultados de clasificación posiblemente podrían encontrarse utilizando electrodos fuera de dicha región de la corteza cerebral, pero realizar esto implica utilizar señales no-específicas o no-correlacionadas con la tarea mental sensorial de interés. De hecho, la configuración “óptima” de canales es aún un tema abierto en el campo de las BCIs. Hay falta de consenso no sólo en el número “óptimo” sino también en las posiciones “óptimas” de los electrodos a utilizar para una aplicación de BCI particular.

Actualmente uno de los principales objetivos es trasladar las BCIs fuera del laboratorio, para que sean utilizadas por sus potenciales usuarios finales en sus vidas cotidianas. Esto requiere de sistemas que sean lo más independientes posibles de la presencia constante de técnicos o profesionales especializados. Asimismo, las herramientas utilizadas para medir la actividad cerebral deben ser no invasivas, transportables, seguras y asequibles. Este último requerimiento limita en gran medida el uso de las BCIs a clínicas especializadas o a laboratorios de investigación. Para romper esta barrera se plantea el desarrollo de una base de datos utilizando equipos de bajo costo y transportables denominados OpenBCI<sup>1</sup> junto con el software de experimentación libre y gratuito OpenVIBE, para así evaluar si es posible construir una BCI confiable mediante estos sistemas. Los resultados preliminares que tenemos hasta el momento son prometedores, y parecen indicar que BCIs de bajo costo pueden efectivamente construirse y utilizarse.

Otros trabajos futuros incluyen el estudio de métodos de inferencia para la estimación

---

<sup>1</sup><http://openbci.com/>

sujeto-específica de los parámetros de regularización tales como los métodos basados en aprendizaje Bayesiano (Li et al., 2010), el principio de “balanceo” (Ito et al., 2011), el método de agregación (Chen et al., 2015), etc. Por otro lado, el enfoque conocido como “LASSO por grupos” (Yuan y Lin, 2006) ha sido muy utilizado para la selección automática sujeto-dependiente de los electrodos más discriminativos (Flamary et al., 2014; Yu et al., 2015). Estas y similares ideas pueden ser incluidas en nuestro modelo. Finalmente, dados los resultados alcanzados por el método PFBCSP en tiempo real en cada ronda de evaluación, es necesario explorar la cantidad mínima de épocas requeridas para detectar eficientemente MI, técnicas adaptativas de entrenamiento, metodologías de transferencia de “sesión-a-sesión”, por nombrar algunos.

Para concluir, consideramos que los principales aportes presentados en este trabajo de tesis han contribuido al avance de técnicas discriminativas que pueden aplicarse más allá del ámbito de las BCIs, con la esperanza de que sean utilizadas por la comunidad. Los resultados aquí expuestos junto con otros trabajos han originado ya seis publicaciones a congresos internacionales y nacionales como así también trabajos en revistas internacionales con referato. A continuación, listadas por su orden de publicación, se encuentran detalladas dichas publicaciones:

- V. Peterson, H.L. Rufiner y R.D. Spies. “Ad-hoc Gaussian Dictionaries for Sparse Representation of Evoked Related Potentials”, *Proceedings of the 1st Pan-American Congress on Computational Mechanics*, Buenos Aires, Argentina, 2015. E-book, ISBN: 978-84-943928-2-5.
- V. Peterson, H.L. Rufiner y R.D. Spies. “L1-Norm Regularization for Sparse Representation and P300 Wave Detection in Brain-Computer Interfaces”, *Proceedings of the V Congreso de Matemática Aplicada, Computacional e Industrial*, Tandil, Argentina, 2015. E-book, ISSN: 2314-3282.
- V. Peterson, H.L. Rufiner y R.D. Spies. “Generalized Sparse Discriminant Analysis for Event-Related Potential Classification”, *Biomedical Signal Processing and Control*, Vol. 35, pp. 70-78, 2017.
- V. Peterson, H.L. Rufiner y R.D. Spies. “Analysis of different discriminant measures on a penalized mix-norm classification method for ERP detection”, *Proceeding of VI Congreso de Matemática Aplicada, Computacional e Industrial (VI MACI 2017)*, Comodoro Rivadavia, Argentina, 2017.
- V. Peterson, H.L. Rufiner y R.D. Spies. “On the measurement impact of the a-priori discriminative information in a penalized classification method for ERP detection”, *Pattern Recognition Letters. En Revisión*.



- V. Peterson, D. Wyser, O. Lambercy, R.D. Spies y R. Gasert. “A penalized time-frequency band feature selection and classification procedure for improved motor intention decoding in multichannel EEG”, *Journal of Neural Engineering*. *En Revisión*.

Asimismo, con la idea de contribuir abiertamente a la comunidad, se ha realizado un tutorial para el uso de OpenBCI junto con OpenVIBE denominado “OpenBCI & OpenVIBE for P300 speller paradigm: a quick starting guide”, publicamente accesible en GitHub <https://github.com/vpeterson/OV-0BCI>. Los códigos de GSDA para Matlab con una pequeña demo también se encuentran disponibles en <https://github.com/vpeterson/gsda>.

# Anexos



# Anexo A

## Detalles sobre FLDA y CSP

### A.1. FLDA como un problema de mínimos cuadrados

Sea  $\mathbf{X}$  una matriz de  $n \times p$ , y sea  $\mathbf{y} \in \mathbb{R}^n$  tal que  $y_i = \frac{n_2}{n}$  ó  $y_i = -\frac{n_1}{n}$ , dependiendo de que la  $i$ -ésima observación pertenezca a la clase 1 o a la clase 2, respectivamente. Consideremos el siguiente problema de regresión de mínimos cuadrados (LSR, por sus siglas en inglés):

$$\min_{\boldsymbol{\alpha} \in \mathbb{R}^p} \frac{1}{2} \|\mathbf{y} - \mathbf{X}\boldsymbol{\alpha}\|_2^2, \quad (\text{A.1})$$

cuyas soluciones son todos los vectores que se encuentra en el conjunto  $\mathcal{N}(\mathbf{X}^T\mathbf{X}) + (\mathbf{X}^T\mathbf{X})^\dagger \mathbf{X}^T \mathbf{y}$ , donde “ $\dagger$ ” denota la inversa generalizada de Moore-Penrose y  $\mathcal{N}(\mathbf{X}^T\mathbf{X})$  denota el espacio nulo de  $\mathbf{X}^T\mathbf{X}$ . Si  $\mathbf{X}^T\mathbf{X}$  es invertible, luego (A.1) tiene solución única dada por  $\boldsymbol{\alpha} = (\mathbf{X}^T\mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}$ . Si la matriz de datos  $\mathbf{X}$  está centrada a media cero ( $\boldsymbol{\mu} = 0$ ), luego  $\mathbf{X}^T\mathbf{X} = n\boldsymbol{\Sigma}_t$ ,  $\mathbf{X}^T \mathbf{y} = \frac{n_1 n_2}{n} (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)$ , y por lo tanto  $\boldsymbol{\alpha} = \frac{n_1 n_2}{n} \boldsymbol{\beta}$ . De esta manera queda demostrado que FLDA también puede ser resuelto mediante un problema de mínimos cuadrados.

### A.2. Método CSP

Sea  $\{\mathbf{X}_c^i\}_{i=1}^{n_c}$  un conjunto de  $n_c$  épocas de EEG filtradas, donde  $c = 1, 2$  representa cada uno de las dos posibles clases. La matriz  $\mathbf{X}_c^i$  de dimensión  $p \times m$ , con  $p$  y  $m$  denotando el número de canales y el número de puntos muestrales, respectivamente, es un época de EEG. En lo que sigue se asume que cada época de EEG tiene media zero y varianza unitaria. El objetivo de CSP es encontrar  $p$  filtros espaciales,  $\mathbf{w}_1, \dots, \mathbf{w}_p \in \mathbb{R}^p$ , tales que

transformen linealmente la matriz  $\mathbf{X}$  según

$$\mathbf{Z} = \mathbf{W}^T \mathbf{X}, \quad (\text{A.2})$$

donde  $\mathbf{W} = [\mathbf{w}_1, \dots, \mathbf{w}_p]$  y  $^T$  denota la transpuesta. El criterio de optimización seguido por CSP para encontrar los  $p$  filtros espaciales requiere de estimaciones muestrales de las matrices de covarianza de cada clase:

$$\Sigma_c = \frac{1}{n_c} \sum_{i=1}^{n_c} \mathbf{X}_c^i (\mathbf{X}_c^i)', \quad c = 1, 2. \quad (\text{A.3})$$

Luego, se procede con la diagonalización simultánea de las dos matrices de covarianza  $\Sigma_1$  y  $\Sigma_2$ :

$$\mathbf{W}^T \Sigma_1 \mathbf{W} = \Delta_1, \quad (\text{A.4})$$

$$\mathbf{W}^T \Sigma_2 \mathbf{W} = \Delta_2, \quad (\text{A.5})$$

donde cada columna de la matriz  $\mathbf{W}$  es escalada de manera que las dos matrices estrictamente positivas  $\Delta_c$ ,  $c = 1, 2$ , satisfagan  $\Delta_1 + \Delta_2 = \mathbf{I}$ . Mediante  $\gamma_j = \frac{\delta_1^j}{\delta_2^j}$ , donde  $\delta_c^j$  es el  $j$ -ésimo elemento diagonal de  $\Delta_c$ , revolver las ecuaciones (A.4) y (A.5) se puede revolver mediante un problema de autovalores generalizado, dado que

$$\mathbf{w}_j^T \Sigma_1 \mathbf{w}_j = \delta_1, \quad (\text{A.6})$$

$$\mathbf{w}_j^T \Sigma_2 \mathbf{w}_j = \delta_2. \quad (\text{A.7})$$

Luego,

$$\Sigma_1 \mathbf{w}_j = \delta_1 \mathbf{w}_j, \quad (\text{A.8})$$

$$\Sigma_1 \mathbf{w}_j = \delta_2 (\delta_1) / \delta_2 \mathbf{w}_j. \quad (\text{A.9})$$

Entonces,

$$\Sigma_1 \mathbf{w}_j = \gamma_j \Sigma_2 \mathbf{w}_j, \quad j = 1, \dots, p. \quad (\text{A.10})$$

## Anexo B

# Ad-hoc Gaussian Dictionaries for Sparse Representation of Evoked Related Potentials

V. Peterson, H.L. Rufiner y R.D. Spies. *Proceedings of the 1st Pan-American Congress on Computational Mechanics*, Buenos Aires, Argentina, 2015. E-book, ISBN: 978-84-943928-2-5.



## AD-HOC GAUSSIAN DICTIONARIES FOR SPARSE REPRESENTATION OF EVOKED RELATED POTENTIALS

VICTORIA PETERSON<sup>1</sup>, HUGO L. RUFINER<sup>2</sup> AND RUBEN D. SPIES<sup>3</sup>

<sup>1</sup>Instituto de Investigación en Señales, Sistemas e Inteligencia Computacional, Facultad de Ingeniería y Ciencias Hídricas, Universidad Nacional del Litoral, Consejo Nacional de Investigaciones Científicas y Técnicas (sinc(i)-FICH-UNL-CONICET).  
Ruta Nac. No 168, km 472.4 (3000), Santa Fe - Argentina  
vpeterson@santafe-conicet.gov.ar

<sup>2</sup>Instituto de Investigación en Señales, Sistemas e Inteligencia Computacional, Facultad de Ingeniería y Ciencias Hídricas, Universidad Nacional del Litoral, Consejo Nacional de Investigaciones Científicas y Técnicas (sinc(i)-FICH-UNL-CONICET)  
Ruta Nac. No 168, km 472.4 (3000), Santa Fe - Argentina and Facultad de Ingeniería, Universidad Nacional de Entre Ríos (FI-UNER) Ruta Prov. No 11, km 10 (3100), Oro Verde - Entre Ríos - Argentina  
lrufiner@fich.unl.edu.ar

<sup>3</sup>Instituto de Matemática Aplicada del Litoral (IMAL-CONICET-UNL)  
Ruta Nac. No 168, Paraje El Pozo (3000), Santa Fe - Argentina and Facultad de Ingeniería Química, Universidad Nacional del Litoral, Consejo Nacional de Investigaciones Científicas y Técnicas (FIQ-UNL)  
Santiago del Estero 2829 (3000), Santa Fe - Argentina.  
rspies@santafe-conicet.gov.ar

**Key words:** Brain-Computer Interface, Evoked Related Potentials, Sparse Representation,  $\ell_1$ -minimization, Robust Classification

**Abstract.** A Brain-Computer Interface (BCI) is a system which provides direct communication between the mind of a person and the outside world by using only brain activity (EEG). A common EEG BCI paradigm is based on the so called Event-Related Potentials (ERP) which are responses of the brain to some external stimuli. For the present work at hand, the innermost part of a BCI is the pattern recognition stage whose aim is to detect the presence of ERPs with high accuracy. In recent years there has been a growing interest in the study of sparse representation of signals. Using a dictionary composed of prototype atoms, signals are written as linear combinations of just a few of those atoms. This sparse representation is found to be appropriate for posterior classification purposes. In this work we propose a sparse representation and posterior classification of ERPs signals by means of an ad-hoc spatio-temporal dictionary composed of bidimensional Gaussian atoms. Based on  $\ell_1$ -minimization we find the sparsest possible solution which allow us to design a robust classification based on nearest representation



## 1 INTRODUCTION

Sparse representations have received great interest in the recent years due to their success in many applications in signal and image processing [1], [2].

Given a  $m \times n$  matrix  $A$ , called dictionary, an unknown signal  $\mathbf{x} \in \mathbb{R}^n$  and a measurement vector  $\mathbf{y} \in \mathbb{R}^m$ , we seek to find the sparsest coefficient vector  $\mathbf{x}$  such that  $\mathbf{y} = A\mathbf{x}$ . Recovering  $\mathbf{x}$  given  $A$  and  $\mathbf{y}$  is a non trivial inversion problem since in general the size of  $\mathbf{x}$  is greater than the size of  $\mathbf{y}$ , the measurement  $\mathbf{y}$  is contaminated by noise and the problem is usually severally ill-posed. Hence, regularization is required.

Given a functional  $J(\mathbf{x})$ , which penalizes certain undesired properties of the solution, a regularized solution can then be obtained by solving the following constrained optimization problem:

$$(P_J) : \min_{\mathbf{x} \in \mathbb{R}^n} J(\mathbf{x}) \quad s.t. \quad \mathbf{y} = A\mathbf{x}. \quad (1)$$

Most of the existing works on sparse learning are based on variants of the  $\ell_1$ -norm regularization ( $J(\mathbf{x}) = \|\mathbf{x}\|_1$ ) due to its sparsity-inducing property, convenient convexity, wide and strong theoretical support, and great success in several applications [3]. Thus, the regularized sparse solution are obtained by solving the following minimization problem:

$$(P_1) : \min_{\mathbf{x} \in \mathbb{R}^n} \|\mathbf{x}\|_1 \quad s.t. \quad \mathbf{y} = A\mathbf{x}. \quad (2)$$

It has been shown, under certain conditions, that the minimum  $\ell_1$ -norm solution of an undetermined linear system is also the sparsest possible solution [4]. In practice, however,  $\mathbf{y}$  is contaminated by noise and therefore the equality constraint in (2) must be relaxed:

$$(P_{1,2}) : \min_{\mathbf{x} \in \mathbb{R}^n} \|\mathbf{x}\|_1 \quad s.t. \quad \|\mathbf{y} - A\mathbf{x}\|_2 \leq \epsilon, \quad (3)$$

where  $\epsilon$  is noise level.

Problem  $(P_{1,2})$ , known as the *basis pursuit denoising* problem (BPDN), is equivalent to the following unconstrained minimization problem:

$$(P_\lambda) : \min_{\mathbf{x} \in \mathbb{R}^n} \frac{1}{2} \|\mathbf{y} - A\mathbf{x}\|_2^2 + \lambda \|\mathbf{x}\|_1, \quad (4)$$

which can be viewed as a generalized Tikhonov-Phillips regularization functional [5] or in a statistical context as a Least Absolute Shrinkage and Selection Operator (LASSO) [6].

In this article we pretend to use the virtues of sparse representation framework together with the ideas proposed in [7] in order to develop a robust classification method in the context of Event Related Potential (ERP) detection.

The problem at hand comes from Brain-Computer Interfaces (BCI) systems. BCI can significantly improve the quality of life of a person who cannot control his/her own body or even is not able to communicate. By using only brain activity, BCI provides a person a new way of communication and control without needing any peripheral nerves or muscles

[8]. The most common and non-invasive method used to decode the intention of a BCI user consists of detecting the presence of ERP signals in electroencephalogram (EEG) records.

In the “oddball” paradigm it is well known that when a person is stimulated with some external and “rare” item (which can be auditory, visual or somatosensorial) an ERP is elicited. One of the main components of ERP signals is an enhanced positive-going component with a latency of about 300 ms (called P300 wave) [9], [10]. Unfortunately detecting a P300 wave (which means detecting the ERP signal) is not an easy task. Mainly due to the fact that SNR between ERP and EEG signals is very low (about -50 db) and also due to the large variation present in P300 wave records among different trails.

In order to use all the information that can be found in EEG records we shall construct an ad-hoc dictionary whose elements are composed by two bidimensional Gaussian atoms representing the main spatio-temporal variation of EEG records with and without P300, respectively. As explained in the next section, the dictionary was built based on a certain neurologically well-supported assumptions on the P300 wave.

## 2 MATERIALS AND METHODS

### 2.1 Database

An Open-Access P300 speller database was used [11]. EEG records from 18 subjects were acquired by 10 electrodes in the positions shown in Fig.1. The records were digitalized at a rate of 256 *Hz*.

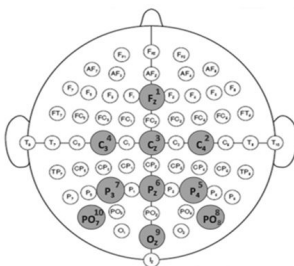


Figure 1: Positions of the electrodes used in EEG acquisition. Modified from [11].

The Farwell and Donchin oddball paradigm was used [10], where a 6-by-6 matrix containing letters and numbers were displayed on a computer screen. During the experiment, a subject is asked to spell different words. The person must focus on one character at the time. As stimulus, a row or a column of the matrix is randomly flashed. In each stimulating block, every row and column of the matrix is intensified only once. If the person is well concentrated, when the chosen character is illuminated, a relevant event occurs, i.e. an ERP signal is elicited [12]. Since there are two classes involved, the 6-by-6 matrix result in twelve possible events (one per each row and one per each column), of which only two are relevant, i.e. with ERP, in every intensification block.

Each subject participated in 4 sessions, where the first two ones were copy-spelling runs. For this reason, in this work we used the first and second sessions as training and testing sets, respectively. Each session consisted of 15 trials, resulting in 2880 EEG epochs per channel (480 target records and 2400 non-target records) for the training set and 900 EEG epochs per channel (150 of them being target) for the testing set. An epoch is a EEG record of one second duration extracted at the beginning of an intensification.

## 2.2 Gaussian Dictionary Generation

### 2.2.1 Image Pre-processing

In order to use all the information available in an epoch, we used not only the information time-to-time but also the spatial information given by the electrode’s positions. In this way, we constructed one image per trial in the time-channel plane. Since a trial is an epoch of one second duration extracted at the beginning of an intensification, and the number of channels was ten, our images consisted of  $256 \times 10$  pixels.

Analyzing the images of the great averages for both classes (target records and non-target records), we noticed that the shape of the P300’s peaks could be significantly improve by re-ordering the channels. In order to do this we proceeded as follows. Since the registered amplitude of the P300 wave differs according to the sensors’s position, the channels were re-ordered based on decreasing signal energy, i.e, we chose as the first channel the one possessing greatest  $\ell_2$ -norm, and the rest were decreasingly ordered by their Euclidean distance at this “first” channel. Next we filtered the image with a median filter. An example of a resulting re-ordered and filtered image are shown in column 2 of both figures Fig.2a and Fig.2b.

After analyzing each re-ordered and filtered target image, we observed that for all of them there were one or two notable peaks between 0.2s and 0.6s. On the other hand, the peaks for non-target images always came from the oscillatory background. For those reasons and because the 0.2-0.6 range is in agreement with the latency window of the P300 wave, we cropped all images between 0.2s and 0.6s, resulting in images of  $104 \times 10$  pixels.

Images belonging to the three different pre-processing stages are shown in Fig.2 for the first subject in the database. There are the original image, the re-ordered and filtered image and the re-ordered, filtered and cropped image for both classes in two different views.

In the sequel, a template will referred to the re-ordered, filtered and cropped images.

### 2.2.2 Levenberg-Marquardt Estimation of Mother Elements

Inspired by the ideas of pre-defined dictionaries (like wavelet dictionaries), we want to generate a dictionary by means of variations of an appropriate “mother” element.

In our case we shall consider one mother element per class and per subject, given by a

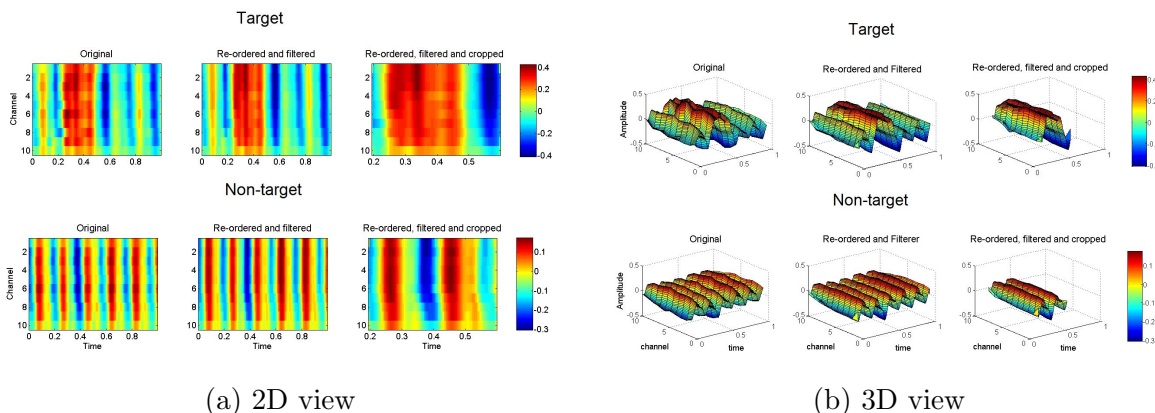


Figure 2: Target and non-target images for subject N° 1. From left to right: original image, re-ordered and filtered image and re-ordered, filtered and cropped image. Fig. 2a show a 2D view while Fig.2b show a 3D view

linear affine combination of two Gaussian functions (atoms). More precisely, let

$\mathcal{P} = \{(p_1, p_2, \dots, p_{11}) \in \mathbb{R}^{11} : p_5, p_6, p_{10}, p_{11} > 0\}$  and for  $\mathbf{p} \in \mathcal{P}$  define:

$$z(t, c; \mathbf{p}) = p_1 + p_2 \exp\left(\frac{(t - p_3)^2}{p_5^2} + \frac{(c - p_4)^2}{p_6^2}\right) + p_7 \exp\left(\frac{(t - p_8)^2}{p_{10}^2} + \frac{(c - p_9)^2}{p_{11}^2}\right). \quad (5)$$

In the sequel we shall always identify the scalar field  $z(t, c; \mathbf{p})$  with the vector  $z(\mathbf{p}) \in \mathbb{R}^{1040}$  obtained after stacking in a column vector the matrix resulting of the evaluation of  $z(t, c; \mathbf{p})$  over  $104 \times 10$  image grid.

Given the template  $f(t, c)$  (or simply  $f \in \mathbb{R}^{1040}$ ) we formulate the corresponding non-linear fitting problem associated to  $f$  and the mother element as:

$$(F_{\mathbf{p}}) : \min_{\mathbf{p} \in \mathcal{P}} \|f - z(\mathbf{p})\|_2^2. \quad (6)$$

The Levenberg-Marquardt (LM) method [13], [14] is a standard and efficient technique to solve nonlinear least squares problems. We used it to find  $\mathbf{p}$  in (6).

The estimated mother elements for the same subject used in Fig.2 are shown in Fig.3. The templates are shown on the left while the two estimated Gaussian elements fitted by LM algorithm are on the right.

It is timely to observe the good fitting of the mother elements to the corresponding templates, especially for the target class.

### 2.2.3 Dictionary Build Up

As the dictionary must capture the main variation of the P300 wave for each subject, in order to construct an appropriate dictionary it is very important to analyze the sensitivity

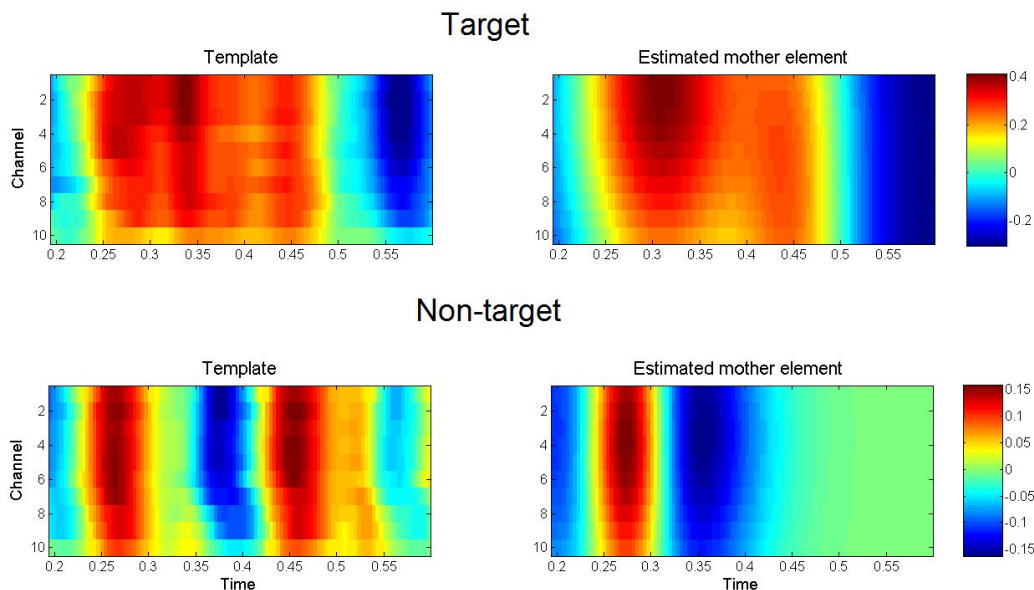


Figure 3: Templates for subject N° 1 (left) and mother elements estimated by LM method (right).

of each one of the components of the parameters vector  $\mathbf{p}$  with respect to the epochs in each one both classes. With this in mind, we performed a rough analysis in this direction following the next steps:

1. **Averaging and balancing:** take averages of 5 trials for target records and averages of 25 trials for non-target records (see Remark (1) below).
2. **Template generation:** for each one of the averaged trials generated in step (1) construct the corresponding template image, as described in Section 2.2.1.
3. **Parameters generation:** for each one of the templates described in step (2) estimate the corresponding parameter  $\mathbf{p}$  by solving problem (6) with the LM method.
4. **Parameters range estimation:** step (3) generates a family of vectors  $\mathbf{p} \in \mathbb{R}^{11}$  for each one of both classes. The range variation for each one of the 11 components of  $\mathbf{p}$  is estimated within each one of both families in the usual way.

**Remark (1).** *The different number of elements, 5 and 25, for taking averages in each one of the classes has the objective of balancing the different number of epochs among classes.*

For the dictionary generation itself we varied one parameter at the time while the others were kept constant. The variations were made in order to cover the whole range

in each one of the components of  $\mathbf{p}$  in one hundred equal increments<sup>1</sup>.

From the sparse representation point of view, it is highly desirable not to have dictionary's elements which are "too similar". The Mutual Coherence (MC) of a dictionary, denoted by  $\mu(A)$ , measures the similarity between dictionary's elements. It is defined as the maximal absolute scalar product between two different  $\ell_2$ -normalized elements of the dictionary  $A$  [1], that is:

$$\mu(A) \doteq \max_{i \neq j} | \langle \mathbf{a}_i, \mathbf{a}_j \rangle |. \quad (7)$$

Notice that if two dictionary's elements are parallel then  $\mu(A) = 1$ , otherwise  $\mu(A) \geq 0$ .

In the dictionary generation algorithm we discarded an element if its MC was greater than some predefined number  $\kappa$ . To avoid classification bias, the dictionary sizes should be the same for both classes. With this objective in mind, we randomly eliminated the necessary number of elements from the larger dictionary whose MCs were greater than another predefined value (e.g.  $\kappa - 0.5$ ).

At this point we have for any subject one dictionary per class, both dictionaries having the same number of elements and with some predefined  $\text{MC}=\kappa$ . It is appropriate to mention here that the size of the dictionary grows with  $\kappa$ .

### 2.3 Classification Based on Sparse Representation

Let us define a new matrix  $A$  as the concatenation of the  $n$  elements from both target and non-target dictionaries,  $A_1$  and  $A_2$ , respectively, that is:

$$A \doteq [A_1 \ A_2] = [\mathbf{a}_{1,1}, \mathbf{a}_{1,2}, \dots, \mathbf{a}_{1,n}, \mathbf{a}_{2,1}, \dots, \mathbf{a}_{2,n}]. \quad (8)$$

For given  $\epsilon$ ,  $A$  and  $\mathbf{y}$ , let  $\hat{\mathbf{x}}$  be the solution of the problem  $(P_{1,2})$  in (3). Ideally one would expect that the nonzero entries of  $\hat{\mathbf{x}}$  will correspond to columns of  $A$  belonging all to the same class. In that case the association of  $\mathbf{y}$  to one of both classes is clearly trivial. Noise and modeling errors, nonetheless, may lead to nonzero entries associated to the wrong class. It is therefore clear that any classification criterion will have to take into account additional information such as the goodness of fit, that is, a measure of how well the coefficients associated with each one of the classes reproduced  $\mathbf{y}$ .

For each class  $i$ , let  $\delta_i : \mathbb{R}^n \rightarrow \mathbb{R}^n$  be the lifting function that selects the coefficients associated with the  $i^{\text{th}}$  class, that for  $\mathbf{x} \in \mathbb{R}^n$ ,  $\delta_i(\mathbf{x}) \in \mathbb{R}^n$  is the vector whose only nonzero entries are the elements in  $\mathbf{x}$  that are associated to class  $i$ . The representation of a given test sample  $\mathbf{y}$  in class  $i$  is then  $\hat{\mathbf{y}}_i = A\delta_i(\hat{\mathbf{x}})$ . The classification of  $\mathbf{y}$  proceeds by assigning it to the class that minimizes the residual, i.e., we associate to  $\mathbf{y}$  the class given by:

$$\underset{i=1,2}{\operatorname{argmin}} r_i(\mathbf{y}) \doteq \|\mathbf{y} - A\delta_i(\hat{\mathbf{x}})\|_2. \quad (9)$$

The Sparse Representation-based Classification (SRC) algorithm (as proposed in [7]) can then be written as follows:

---

<sup>1</sup>The variation in  $p_1$ , corresponding to the offset parameter in (5) was negligible and therefore it was kept constant for the dictionary generation.

1. **Input:** The train dictionary  $A = [A_1, A_2] \in \mathbb{R}^{m \times n}$ , a test sample  $\mathbf{y} \in \mathbb{R}^m$  and  $\epsilon$  tolerance or  $\lambda$  regularized parameter.
2. Normalize the columns of  $A$  to have unit  $\ell_2$  - norm.
3. Solve the  $\ell_1$ -minimization problem:

$$\min_{\mathbf{x} \in \mathbb{R}^n} \|\mathbf{x}\|_1 \quad s.t. \quad \|\mathbf{y} - A\mathbf{x}\|_2 \leq \epsilon.$$

Or equivalently solve:

$$\min_{\mathbf{x} \in \mathbb{R}^n} \frac{1}{2} \|\mathbf{y} - A\mathbf{x}\|_2^2 + \lambda \|\mathbf{x}\|_1.$$

4. Compute the residuals  $r_i(\mathbf{y}) = \|\mathbf{y} - A\delta_i(\hat{\mathbf{x}})\|_2$ , for  $i=1,2$ .
5. **Output:** Identity( $\mathbf{y}$ )= $\underset{i=1,2}{\operatorname{argmin}} r_i(\mathbf{y})$ .

### 3 RESULTS AND DISCUSSIONS

The sparse vectors  $\hat{\mathbf{x}}$  for each test observation  $\mathbf{y}$  were estimated using different functions by the application of SLEP 4.1 toolbox [15]. More precisely we estimated the sparse vectors  $\hat{\mathbf{x}}$  as follows:

1.  $\ell_1$ -ball constrained least squares problem (LeastC):

$$\min_{\mathbf{x}} \frac{1}{2} \|\mathbf{y} - A\mathbf{x}\|_2^2 \quad s.t. \quad \|\mathbf{x}\|_1 < \epsilon, \quad (10)$$

where  $\epsilon$  is the radius of the  $\ell_1$ -ball.

2.  $\ell_1$ -norm regularized least squares problem (LeastR):

$$\min_{\mathbf{x}} \frac{1}{2} \|\mathbf{y} - A\mathbf{x}\|_2^2 + \lambda \|\mathbf{x}\|_1. \quad (11)$$

3. Non-negative  $\ell_1$ -ball constrained least squares problem (NNLeastC):

$$\min_{\mathbf{x}} \frac{1}{2} \|\mathbf{y} - A\mathbf{x}\|_2^2 \quad s.t. \quad \|\mathbf{x}\|_1 < \epsilon, \mathbf{x} \geq 0. \quad (12)$$

4. Non-negative  $\ell_1$ -norm regularized least squares problem (NNLeastR):

$$\min_{\mathbf{x} \geq 0} \frac{1}{2} \|\mathbf{y} - A\mathbf{x}\|_2^2 + \lambda \|\mathbf{x}\|_1. \quad (13)$$

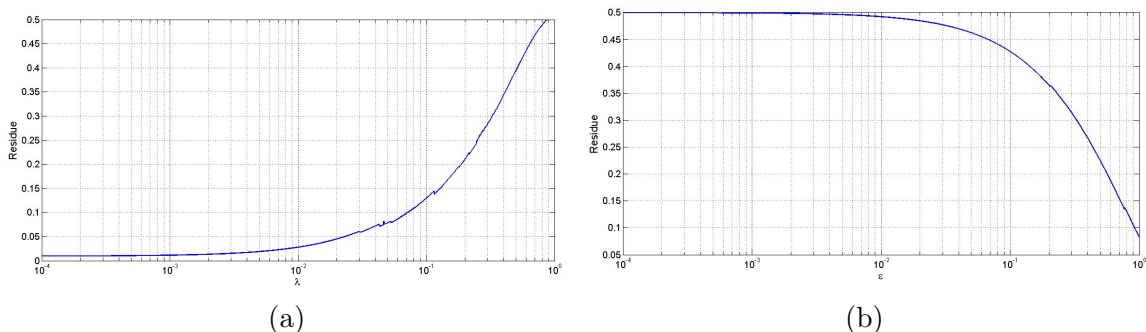


Figure 4: Residue as a function of  $\lambda$  for the LeastR algorithm (4a) and residue as a function of  $\epsilon$  for the LeastC algorithm (4b)

The optimal values of  $\lambda$  and  $\epsilon$  were fixed after analyzing the residue as a function of  $\lambda$  or  $\epsilon$ , respectively. We varied the parameters between  $10^{-4}$  and 1 in increment of  $10^{-4}$ . We finally chose  $\lambda = 0.05$  and  $\epsilon = 0.8$  as optimal values. The behavior of the residue is shown in Fig.4.

In order to analyze the impact of the choice of MC value on generation of the dictionary and, consequently, on the classification rate obtained by SRC algorithm, we varied the MC values from 0.500 to 0.995 in increments of 0.005.

The unweighted accuracy rate (UAR) was used as the performance classification measure. Fig.5 shows the classification results reached by each optimization problem (LeastC, LeastR, NNLeastC and NNLeastR) for four different database's subjects.

There does not seem to be a clear tendency between classification results and MC values. Note that while for subject N° 10 the UAR shows a clearly increasing tendency for increasing MC values, the opposite happens for subject N° 1. This observation constitutes a solid reason for developing a subject depended classification tools.

Table 1 summarizes the best UAR results per each subject with the different optimization problems. The corresponding MC value is shown between parenthesis. An analysis of Table 1 seems to suggest that better classification result are obtained by imposing  $\mathbf{x} \geq 0$ . Although further analysis is required, we strongly believe that this is due to the fact that the elements in the target dictionary present mainly positive peaks.

## 4 CONCLUSIONS

In this paper, we constructed an ad-hoc Gaussian dictionary for representing the P300 wave in a channel-time space. We focused our work in constructing a suitable dictionary per each subject in order to represent the variation of the P300 wave of that particular subject. It is not of our interest to find a generalized (all subjects) representation of the P300 wave. Moreover, we have good reasons to believe that better classification result can be obtained by improving the representation of non-target signals.



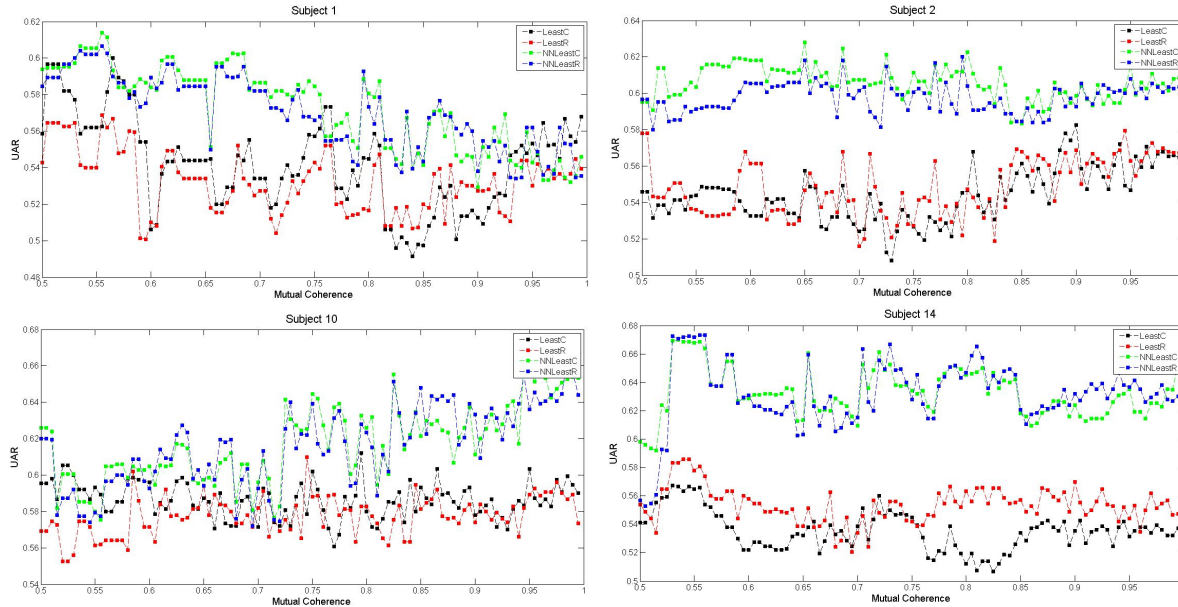


Figure 5: UAR results as a function of MC values for the four different algorithms used.

The resulting  $\ell_1$ -minimization problems were solved using two different optimization approach both with and without a positive constraint, and used the sparse information of the  $\hat{\mathbf{x}}$  solution for classification purposes. Although the classification performances are far from being optimal, we point out that the classification method used is very simple and moreover noisy single trail epochs were used. Much further research is clearly needed in this regard. In particular it is of great interest to find alternative and/or complementary ways to the  $\ell_1$ -minimization approach, which could allow classification improvement. We are currently devoting efforts in this direction.

## 5 ACKNOWLEDGEMENTS

This work was supported in part by Consejo Nacional de Investigaciones Científicas y Técnicas, CONICET, through PIP 2014-2016 Nro. 11220130100216-CO, the Air Force Office of Scientific Research, AFOSR/SOARD, through Grant FA9550-14-1-0130 and by CAID-UNL 2011.

## REFERENCES

- [1] Elad, M. Sparse and Redundant Representation, from theory to applications in signal and image processing. *Springer*. (2010). ISBN 978-1-4419-7010-7.
- [2] Bruckstein, J.M, Donoho, D. and Elad, M. From Sparse Solutions of Systems of Equations to Sparse Modeling of Signals and Images. *Society for Industrial and Applied*

Table 1: Best UAR results for the different optimization problems.

Subject N°	LeastC	LeastR	NNLeastC	NNLeastR
1	0.600 (0.565)	0.569 (0.555)	0.614 (0.555)	0.607 (0.555)
2	0.583 (0.900)	0.579 (0.945)	0.628 (0.650)	0.620 (0.795)
3	0.531 (0.680)	0.531 (0.730)	0.503 (0.780)	0.509 (0.680)
4	0.558 (0.970)	0.546 (0.645)	0.583 (0.955)	0.590 (0.965)
5	0.528 (0.670)	0.540 (0.700)	0.559 (0.570)	0.567 (0.600)
6	0.547 (0.855)	0.533 (0.750)	0.539 (0.855)	0.538 (0.855)
7	0.554 (0.645)	0.558 (0.900)	0.545 (0.980)	0.548 (0.995)
8	0.576 (0.920)	0.561 (0.790)	0.537 (0.995)	0.539 (0.995)
9	0.531 (0.650)	0.544 (0.595)	0.551 (0.700)	0.604 (0.635)
10	0.612 (0.795)	0.610 (0.745)	0.663 (0.960)	0.664 (0.985)
11	0.557 (0.910)	0.565 (0.940)	0.595 (0.995)	0.624 (0.940)
12	0.519 (0.680)	0.517 (0.600)	0.545 (0.635)	0.547 (0.640)
13	0.580 (0.865)	0.577 (0.840)	0.639 (0.740)	0.647 (0.625)
14	0.567 (0.530)	0.586 (0.540)	0.670 (0.535)	0.673 (0.555)
15	0.573 (0.700)	0.561 (0.645)	0.585 (0.960)	0.559 (0.855)
16	0.523 (0.585)	0.523 (0.590)	0.614 (0.995)	0.609 (0.995)
17	0.577 (0.590)	0.581 (0.855)	0.589 (0.995)	0.595 (0.605)
18	0.546 (0.920)	0.545 (0.760)	0.577 (0.975)	0.577 (0.980)

*Mathematics*. (2009) **51**:34–81.

- [3] Donoho, D. For Most Large Underdetermined Systems of Equations, the Minimal  $l_1$ -norm Near-Solution Approximates the Sparsest Near-Solution. *Communications on Pure and Applied Mathematics*. (2006) **59**:907–934.
- [4] Donoho, D. and Elad, M. Optimally sparse representation in general (nonorthogonal) dictionaries via  $l_1$  minimization. *Proceeding of the National Academy of Sciences of the United States of America*. (2003) **100**:2197–2202.
- [5] Engl, H.W., Hanke, M. and Neubauer, A. Regularization of inverse problems. *Kluwer Academic Publishers*. (2010). ISBN 0-7923-4157-0.
- [6] Tibshirani, R. Regression, Shrinkage and Selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*. (1996) **58**:267–288.
- [7] Wright, J., Yang, A.Y., Ganesh, A., Sastry, S.S. and Ma, Y. Robust Face Recognition via Sparse Representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. (1996) **31**:1–18.

- [8] Wolpaw J.R., Birbaumer N., McFarland D.J., Pfurtscheller G., Vaughan T.M. Brain computer interfaces for communication and control. *Clinical Neurophysiology*. (2002) **113**:767–791.
- [9] Hiyard S.A., Kutas M. Electrophysiology of cognitive processing *Annual Reviews Psychol.* (1983) **34**:33–61.
- [10] Farwell L.A., Donchin E. Talking off the top of your head: toward a metal prosthesis utilizing event-related brain potentials *Electroencephalography and clinical neurophysiology*. (1988) **70**:510–523.
- [11] Ledesma-Ramirez C., Bojorges-Valdez E., Yañez-Suarez O., Saavedra C., Boygrain L., and Gentiletti G. An Open-Access P300 Speller Database. *Fourth international BCI meeting, Monterrey, USA, California* (2010).
- [12] Donchin E., Spencer K.M. and Wijesinghe R. The mental prosthesis: Assessing the Speed of a P300-Based Brain-Computer Interface. *Transactions on Rehabilitation Engineering*. (2000)**8**:174–179.
- [13] Levenberg, K. A method for the solution of certain problems in least squares. *Quart. Appl. Math.* (1944)**2**:164-168.
- [14] Marquardt, D. An algorithm for least-squares estimation of nonlinear parameters *SIAM J. Appl. Math.* (1963)**11**:431-441.
- [15] Liu J., Ji S. and Ye J. SLEP: Sparse Learning with Efficient Projections. *Arizona State University* (2009) <http://www.public.asu.edu/~jye02/Software/SLEP>.

## Anexo C

# $\ell_1$ -norm Regularization for Sparse Representation and P300 Wave Detection in Brain-Computer Interfaces

V. Peterson, H.L. Rufiner y R.D. Spies. *Proceedings of the V Congreso de Matemática Aplicada, Computacional e Industrial*, Tandil, Argentina, 2015. E-book, ISSN: 2314-3282.



# $\ell_1$ -NORM REGULARIZATION FOR SPARSE REPRESENTATION AND P300 WAVE DETECTION IN BRAIN-COMPUTER INTERFACES

Victoria Peterson<sup>b</sup>, Hugo L. Rufiner<sup>b,†</sup> and Rubén D. Spies<sup>\*</sup>

<sup>b</sup>*Instituto de Investigación en Señales, Sistemas e Inteligencia Computacional, Facultad de Ingeniería y Ciencias Hídricas, Universidad Nacional del Litoral, Consejo Nacional de Investigaciones Científicas y Técnicas (sinc(i)-FICH-UNL-CONICET), Ruta Nac. 168, km 472.4 (3000), Santa Fe - Argentina,*

[vpeterson@santafe-conicet.gov.ar](mailto:vpeterson@santafe-conicet.gov.ar)

<sup>†</sup>*Facultad de Ingeniería, Universidad Nacional de Entre Ríos (FI-UNER), Ruta Prov. 11, km 10 (3100), Oro Verde - Entre Ríos - Argentina, [lrufiner@fich.unl.edu.ar](mailto:lrufiner@fich.unl.edu.ar)*

<sup>\*</sup>*Instituto de Matemática Aplicada del Litoral (IMAL-CONICET-UNL), Ruta Nac. 168, Paraje El Pozo (3000), Santa Fe - Argentina and Facultad de Ingeniería Química, Universidad Nacional del Litoral, Consejo Nacional de Investigaciones Científicas y Técnicas (FIQ-UNL), Santiago del Estero 2829 (3000), Santa Fe - Argentina, [rspies@santafe-conicet.gov.ar](mailto:rspies@santafe-conicet.gov.ar)*

**Abstract:** A Brain-Computer Interface (BCI) is a system which provides direct communication between the mind of a person and the outside world by using only brain activity (EEG). A common EEG-BCI paradigm is based on the so called Event-Related Potentials (ERP) which are responses of the brain to some external stimuli. One of the main components of ERP signals is an enhanced positive-going component called P300 wave. The  $\ell_1$ -norm minimization has been widely used due to its sparsity-inducing property, convenient convexity and great success in several applications. In this work we propose a sparse representation and posterior classification of ERPs signals by means of an ad-hoc spatio-temporal dictionary composed of bidimensional Gaussian elements. The classification is based on minimizing the residual between a test sample and its estimation.

**Keywords:** *Brain-Computer Interface, Sparse Representation,  $\ell_1$ -minimization.*

2000 AMS Subject Classification: primary: 92C55, 92C20, secondary: 65F22, 65J20

## 1 INTRODUCTION

Sparse representations have received great interest in recent years due to their success in many applications in signal and image processing [1], [2]. Given a  $m \times n$  matrix  $A$ , called dictionary, an unknown signal  $\mathbf{x} \in \mathbb{R}^n$  and a measurement vector  $\mathbf{y} \in \mathbb{R}^m$ , we seek to find the sparsest coefficient vector  $\mathbf{x}$  such that  $\mathbf{y} = A\mathbf{x}$ . Recovering  $\mathbf{x}$  given  $A$  and  $\mathbf{y}$  is a non trivial inversion problem since in general the size of  $\mathbf{x}$  is greater than the size of  $\mathbf{y}$ , the measurement  $\mathbf{y}$  is contaminated by noise and the problem is usually severally ill-posed. Hence, regularization is required.

Most of the existing works on sparse learning are based on variants of the  $\ell_1$ -norm regularization due to its sparsity-inducing property, convenient convexity, wide and strong theoretical support, and great success in several applications [3]. In practice  $\mathbf{y}$  is contaminated by noise, thus the regularized sparse solution is obtained by solving the following minimization problem:

$$(P_{1,2}) : \min_{\mathbf{x} \in \mathbb{R}^n} \|\mathbf{x}\|_1 \quad s.t. \quad \|\mathbf{y} - A\mathbf{x}\|_2 \leq \epsilon, \quad (1)$$

where  $\epsilon$  is a noise level.

Problem  $(P_{1,2})$ , known as the *basis pursuit denoising* problem (BPDN), is equivalent to the following unconstrained minimization problem:

$$(P_\lambda) : \min_{\mathbf{x} \in \mathbb{R}^n} \frac{1}{2} \|\mathbf{y} - A\mathbf{x}\|_2^2 + \lambda \|\mathbf{x}\|_1, \quad (2)$$

which can be viewed as a generalized Tikhonov-Phillips regularization method [4] or in a statistical context as a Least Absolute Shrinkage and Selection Operator (LASSO) [5].

In this article we use the sparse representation for posterior classification based on the ideas proposed in [6] in the context of Event Related Potential (ERP) recognition.

The Brain-Computer Interfaces (BCI) can significantly improve the quality of life of a person who cannot control his/her own body or even is not able to communicate. By using only brain activity, BCI provides a person a new way of communication and control without needing any peripheral nerves or muscles [7]. When a person is stimulated with some external and “rare” item (which can be auditory, visual or somatosensory) an ERP is elicited. One of the main components of such ERP signals is an enhanced positive-going component with a latency of about 300 ms (called P300 wave) [8], [9]. Unfortunately, detecting a P300 wave (which means detecting the ERP signal) is not an easy task, mainly due to the fact that SNR between ERP and EEG signals is very low (about -50 db) and also due to the large variation present in P300 wave records among different trails.

In order to use all the information that can be found in EEG records we shall construct an ad-hoc dictionary whose elements are composed by two bidimensional Gaussian atoms representing the main spatio-temporal variation of EEG records with and without P300.

## 2 MATERIAL AND METHODS

### 2.1 DATABASE

An Open-Access P300 speller database was used [10]. EEG records from 3 subjects were acquired by 10 electrodes at a rate of 256  $Hz$ . Each subject participated in 4 sessions. In this work we used the first and second sessions as training and testing sets, respectively. In both sets there are records belonging to the target class (with ERP signals) and belonging to the non-target class (without ERP signals). The training set consisted of 2880 EEG epochs per channel (480 target records and 2400 non-target records) while the testing set consisted of 900 EEG epochs per channel (150 of them being target). In this context, an epoch is a EEG record of one second duration extracted at the beginning of a stimulus.

### 2.2 AD-HOC DICTIONARY GENERATION

In order to use all the information available in an epoch, we used not only the information time-to-time but also the spatial information given by the electrode’s positions. In this way, we constructed one image per trial in the time-channel plane. Because the trial size was 256 and the number of channels was ten, our images consisted of  $256 \times 10$  pixels.

Since the registered amplitude of the P300 wave differs according to the sensors’s position, the channels were re-ordered based on decreasing signal energy. Next we filtered the images with a median filter. Due to the fact that each re-ordered and filtered target image contained one or two notable peaks between 0.2s and 0.6s and because the 0.2-0.6 range is in agreement with the latency window of the P300 wave, we cropped all images between 0.2s and 0.6s, resulting in images of  $104 \times 10$  pixels. In the sequel, a template will be referred to the re-ordered, filtered and cropped image.

Next, we generated a dictionary by means of variations of an appropriate “mother” element. In our case we shall consider one mother element per class and per subject, given by a linear affine combination of two Gaussian functions (atoms). More precisely, let  $\mathcal{P} = \{(p_1, p_2, \dots, p_{11}) \in \mathbb{R}^{11} : p_5, p_6, p_{10}, p_{11} > 0\}$  and for  $\mathbf{p} \in \mathcal{P}$  define:

$$z(t, c; \mathbf{p}) = p_1 + p_2 \exp\left(\frac{(t - p_3)^2}{p_5^2} + \frac{(c - p_4)^2}{p_6^2}\right) + p_7 \exp\left(\frac{(t - p_8)^2}{p_{10}^2} + \frac{(c - p_9)^2}{p_{11}^2}\right). \quad (3)$$

In the sequel we shall always identify the scalar field  $z(t, c; \mathbf{p})$  with the vector  $z(\mathbf{p}) \in \mathbb{R}^{1040}$  obtained after stacking in a column vector the matrix resulting of the evaluation of  $z(t, c; \mathbf{p})$  over  $104 \times 10$  image grid.

Given the template  $f(t, c)$  (or simply  $f \in \mathbb{R}^{1040}$ ) we formulate the corresponding non-linear fitting problem associated to  $f$  and the mother element  $z(\mathbf{p})$  as:

$$(F_{\mathbf{p}}) : \min_{\mathbf{p} \in \mathcal{P}} \|f - z(\mathbf{p})\|_2^2. \quad (4)$$

The Levenberg-Marquardt (LM) method [11], [12], which is a standard and efficient technique to solve nonlinear least squares problems, was used to find  $\mathbf{p}$  in (4).

As the dictionary must capture the main variation of the P300 wave for each subject, in order to construct an appropriate dictionary we analyzed the sensitivity of the parameter vector  $\mathbf{p}$  in each one of both classes. This process was done by solving (4) with templates generated as the average of  $n$  epoch (we took  $n = 5$  for the target class and  $n = 25$  for the non-target class). At the end, we obtained a range variation for each one of the 11 components of  $\mathbf{p}$ . For the dictionary generation itself we varied one parameter at the time while the others were kept constant. The variations were made in order to cover the whole range in each one of the components of  $\mathbf{p}$  in one hundred equal increments<sup>1</sup>.

From the sparse representation point of view, it is highly desirable not to have dictionary elements which are “too similar”. The Mutual Coherence (MC) of a dictionary  $A$ , denoted by  $\mu(A)$ , measures the similarity between dictionary elements. It is defined as the maximal absolute scalar product between two different  $\ell_2$ -normalized elements of the dictionary  $A$  [1]. In the dictionary generation algorithm we discarded an element if its MC was greater than some predefined number  $\kappa$ . To avoid classification bias, the dictionary sizes were required to be the same for both classes. With this objective in mind, we randomly eliminated the necessary number of elements from the larger dictionary whose MCs were greater than another predefined value (e.g.  $\kappa = 0.5$ ).

### 2.3 CLASSIFICATION BASED ON SPARSE REPRESENTATION

Let us define a new matrix  $A$  as the concatenation of the  $n$  elements from both target and non-target dictionaries,  $A_1$  and  $A_2$ , respectively, that is:

$$A \doteq [A_1 \ A_2] = [\mathbf{a}_{1,1}, \mathbf{a}_{1,2}, \dots, \mathbf{a}_{1,n}, \mathbf{a}_{2,1}, \dots, \mathbf{a}_{2,n}]. \quad (5)$$

For given  $\epsilon$ ,  $A$  and  $\mathbf{y}$ , let  $\hat{\mathbf{x}}$  be the solution of the problem  $(P_{1,2})$  in (1). For each class  $i$ , let  $\delta_i : \mathbb{R}^n \rightarrow \mathbb{R}^n$  be the lifting function that selects the coefficients associated with the  $i^{\text{th}}$  class, that for  $\mathbf{x} \in \mathbb{R}^n$ ,  $\delta_i(\mathbf{x}) \in \mathbb{R}^n$  is a vector whose only nonzero entries are the elements in  $\mathbf{x}$  that are associated to class  $i$ . The representation of a given test sample  $\mathbf{y}$  in class  $i$  is then  $\hat{\mathbf{y}}_i = A\delta_i(\hat{\mathbf{x}})$  [6]. The classification of  $\mathbf{y}$  proceeds by assigning it to the class that minimizes the residual, i.e., we associate to  $\mathbf{y}$  the class given by:

$$\operatorname{argmin}_{i=1,2} r_i(\mathbf{y}) \doteq \|\mathbf{y} - A\delta_i(\hat{\mathbf{x}})\|_2. \quad (6)$$

The unweighted accuracy rate (UAR) was used as the performance classification measure. For a binary classification problem, this index is defined as:  $UAR = \frac{1}{2} \left( \frac{TP}{P} + \frac{TN}{N} \right)$ , where  $TP$  and  $TN$  are the number of *true positive* and *true negative*, while  $P$  and  $N$  indicate the amount of samples in each class, respectively.

## 3 RESULTS

The solution of problem  $(P_\lambda)$  was obtained by the application of the SLEP 4.1 toolbox [13]. More precisely, we solved the  $\ell_1$ -norm regularized least squares problem (LeastR), i.e. (2), and the non-negative  $\ell_1$ -norm regularized least squares problem (NNLeastR), where the latter imposes the additional non-negative constrain  $\mathbf{x} \geq 0$  in (2).

The training set was used to fix the “optimal” value of  $\lambda$ . Per each observation  $\mathbf{y}$  we varied  $\lambda$  from 0 to  $10^{-6}$  in a log-scale and analyzed the residue for  $\lambda$  in this range. Minimizing the residual and the sparsity are both desired properties, thus we fixed the range of nonzero entries in  $\mathbf{x}$  between 1 and 10 and defined a new  $\lambda$  range in agreement with this nonzero entries range. We finally chose the  $\lambda$  value that produced the best classification result for the training set and used it for the posterior classification with the testing set.

In order to analyze the impact of the choice of the MC value on the generation of the dictionary and, consequently, on the classification performance, we varied the MC values from 0.90 to 0.95 in increments of 0.005.

Table 1 summarizes the best UAR results for each subject with the different optimization problems. The corresponding MC value is shown between parenthesis. An analysis of Table 1 seems to suggest that better

<sup>1</sup>The variation in  $p_1$ , corresponding to the offset parameter in (3) was negligible and therefore it was kept constant for the dictionary generation.



classification result are obtained by imposing  $\mathbf{x} \geq 0$ . Although further analysis is required, we strongly believe that this is due to the fact that the elements in the target dictionary present mainly positive peaks.

Table 1: Best UAR results for the different optimization problems.

Subject N <sup>o</sup>	LeastR	NNLeastR
1	0.567 (0.985)	0.627 (0.935)
2	0.550 (0.965)	0.576 (0.905)
3	0.495 (0.960)	0.499 (0.955)

## 4 CONCLUSIONS

In the present work we explored the sparse properties for classification purposes of ERP signals. Although the classification performances are far from being optimal, we find it important to point out that the classification method used is very simple and moreover noisy single trial epochs were used. We are currently devoting efforts in this direction. In particular it is of great interest to find alternative and/or complementary ways to the  $\ell_1$ -minimization approach, which could allow classification improvement.

We constructed an ad-hoc Gaussian dictionary for representing the P300 wave in a channel-time space. We focused our work in constructing a subject-dependent dictionary in order to represent the variation of the P300 wave of that particular subject. We have good reasons to believe that better classification results could be obtained by improving the representation of non-target signals.

## ACKNOWLEDGMENTS

This work was supported in part by Consejo Nacional de Investigaciones Científicas y Técnicas, CONICET, through PIP 2014-2016 Nro. 11220130100216-CO, the Air Force Office of Scientific Research, AFOSR/SOARD, through Grant FA9550-14-1-0130 and by CAID-UNL 2011.

## REFERENCES

- [1] M. ELAD. *Sparse and Redundant Representation, from theory to applications in signal and image processing*. Springer, 2010, ISBN 978-1-4419-7010-7.
- [2] J.M. BRUCKSTEIN, D. DONOHO, AND M. ELAD. *From Sparse Solutions of Systems of Equations to Sparse Modeling of Signals and Images*. Society for Industrial and Applied Mathematics, Vol. 51 (2009), pp.34-81.
- [3] D. DONOHO. *For Most Large Underdetermined Systems of Equations the Minimal  $\ell_1$ -norm Near-Solution Approximates the Sparsest Near-Solution*. Communications on Pure and Applied Mathematics, Vol. 59 (2006), pp.907-934.
- [4] H.W ENGL, M. HANKE AND A. NEUBAUER. *Regularization of inverse problems*. Kluwer Academic Publishers, 2010, ISBN 0-7923-4157-0.
- [5] R. TIBSHIRANI. *Regression, Shrinkage and Selection via the lasso*. Journal of the Royal Statistical Society. Series B (Methodological), Vol. 58 (1996), pp.267-288.
- [6] J. WRIGHT, A.Y YANG, A. GANESH, S.S SASTRY AND Y. MA. *Robust Face Recognition via Sparse Representation*. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 31 (1996), pp.1-18.
- [7] J.R WOLPAW, N. BIRBAUMER, D.J MCFARLAND, G. PFURTSCHELLER AND T.M. VAUGHAN. *Brain computer interfaces for communication and control*. Clinical Neurophysiology, Vol. 113 (2002), pp.767-791.
- [8] S.A HIYARD AND M. KUTAS. *Electrophysiology of cognitive processing*. Annual Reviews Psychol, Vol. 34 (1983), pp.33-61.
- [9] L.A FARWELL AND E. DONCHIN, Talking off the top of your head: toward a metal prosthesis utilizing event-related brain potentials *Electroencephalography and clinical neurophysiology*. Vol.70 (1988), pp.510-523.
- [10] C. LEDESMA-RAMIREZ, E. BOJORGES-VALDEZ, O. YAÑEZ-SUAREZ, C. SAAVEDRA, L. BOYGRAIN AND G. GENTILETTI. *An Open-Access P300 Speller Database*. Fourth international BCI meeting, Monterrey, USA, California, 2010.
- [11] K. LEVENBERG. *A method for the solution of certain problems in least squares*. Quart. Appl. Math., Vol. 2 (1944), pp.164-168.
- [12] D. MARQUARDT. *An algorithm for least-squares estimation of nonlinear parameters*. SIAM J. Appl. Math., Vol. 11 (1963), pp.431-441.
- [13] J. LIU, S. JI AND J. YE. *SLEP: Sparse Learning with Efficient Projections*. Arizona State University, 2009. <http://www.public.asu.edu/~jye02/Software/SLEP>.

## Bibliografía

- Allison, Brendan, Dunne, Stephen, Leeb, Robert, Millán, José Del R, y Nijholt, Anton. *Towards practical brain-computer interfaces: bridging the gap from research to real-world applications*. Springer Science & Business Media, 2012.
- Ang, Kai Keng, Chin, Zheng Yang, Wang, Chuanchu, Guan, Cuntai, y Zhang, Haihong. Filter bank common spatial pattern algorithm on BCI competition IV datasets 2a and 2b. *Frontiers in neuroscience*, 6:39, 2012a.
- Ang, Kai Keng, Chin, Zheng Yang, Zhang, Haihong, y Guan, Cuntai. Filter bank common spatial pattern (FBCSP) in brain-computer interface. En *2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence)*, págs. 2390–2397. IEEE, 2008.
- Ang, Kai Keng, Chin, Zheng Yang, Zhang, Haihong, y Guan, Cuntai. Mutual information-based selection of optimal spatial-temporal patterns for single-trial EEG-based BCIs. *Pattern Recognition*, 45(6):2137–2144, 2012b.
- Basseville, Michele. Distance measures for signal processing and pattern recognition. *Signal Processing*, 18:349–369, 1989.
- Belge, Murat, Kilmer, Misha E, y Miller, Eric L. Simultaneous multiple regularization parameter selection by means of the L-hypersurface with applications to linear inverse problems posed in the wavelet transform domain. En *SPIE's International Symposium on Optical Science, Engineering, and Instrumentation*, págs. 328–336. International Society for Optics and Photonics, 1998.
- Benjamin, Emelia J, Blaha, Michael J, Chiuve, Stephanie E, Cushman, Mary, Das, Sandeep R, Deo, Rajat, Floyd, J, Fornage, M, Gillespie, C, Isasi, CR, et al. Heart disease and stroke statistics-2017 update: a report from the American Heart Association. *Circulation*, 135(10):e146–e603, 2017.

- Blankertz, Benjamin, Lemm, Steven, Treder, Matthias, Hauf, Stefan, y Müller, Klaus Robert. Single-trial analysis and classification of erp component- a tutorial. *Neuroimage*, 56:814–825, 2011.
- Blankertz, Benjamin, Muller, K-R, Krusienski, Dean J, Schalk, Gerwin, Wolpaw, Jonathan R, Schlogl, Alois, Pfurtscheller, Gert, Millan, Jd R, Schroder, Michael, y Birbaumer, Niels. The BCI competition III: Validating alternative approaches to actual BCI problems. *IEEE transactions on neural systems and rehabilitation engineering*, 14(2):153–159, 2006.
- Blankertz, Benjamin, Tomioka, Ryota, Lemm, Steven, Kawanabe, Motoaki, y Muller, K-R. Optimizing spatial filters for robust EEG single-trial analysis. *IEEE Signal processing magazine*, 25(1):41–56, 2008.
- Bradley, Andrew P. The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern recognition*, 30(7):1145–1159, 1997.
- Broetz, Doris, Braun, Christoph, Weber, Cornelia, Soekadar, Surjo R, Caria, Andrea, y Birbaumer, Niels. Combination of brain-computer interface training and goal-directed physical therapy in chronic stroke: a case report. *Neurorehabilitation and Neural Repair*, 24(7):674–679, 2010.
- Chen, Jieyang, Pereverzyev Jr, Sergiy, y Xu, Yuesheng. Aggregation of regularized solutions from multiple observation models. *Inverse Problems*, 31(7):075005, 2015.
- Clemmensen, Line, Hastie, Trevor, Witten, Daniela, y Ersbøll, Bjarne. Sparse discriminant analysis. *Technometrics*, 53:406–413, 2012.
- Clerc, Maureen, Bougrain, Laurent, y Lotte, Fabien. *Brain–Computer Interfaces 1: Foundations and Methods*. John Wiley & Sons, Inc., 2016.
- Combrisson, Etienne y Jerbi, Karim. Exceeding chance level by chance: The caveat of theoretical chance levels in brain signal classification and statistical assessment of decoding accuracy. *Journal of neuroscience methods*, 250:126–136, 2015.
- Costa, Sueli IR, Santos, Sandra A, y Strapasson, João E. Fisher information distance: a geometrical reading. *Discrete Applied Mathematics*, 197:59–69, 2015.

- Courty, Nicolas, Flamary, Rémi, y Tuia, Devis. Domain adaptation with regularized optimal transport. En *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, págs. 274–289. Springer, 2014.
- Dickhaus, Thorsten, Sannelli, Claudia, Müller, Klaus-Robert, Curio, Gabriel, y Blankertz, Benjamin. Predicting BCI performance to study BCI illiteracy. *BMC Neuroscience*, 10(1):P84, 2009.
- Donchin, Emanuel, Kevin M., Spencer, y Ranjith, Wijesinghe. The mental prosthesis: Assessing the speed of a P300-Based Brain-Computer Interface. *Transactions on Rehabilitation Engineering*, 8:174–179, 2000.
- Duda, Richard O, Hart, Peter E, y Stork, David G. *Pattern classification*. John Wiley & Sons, 2012.
- Fabiani, Monica, Gratton, Gabriele, Karis, Demetrios, y Donchin, Emanuel. Definition, identification and reability of measurement of the P300 component of the event-related brain potential. *Advances in Psychophysiology*, 2:1–78, 1987.
- Farwell, Lawrence Ashley y Donchin, Emanuel. Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalography and clinical Neurophysiology*, 70(6):510–523, 1988.
- Fawcett, Tom. An introduction to ROC analysis. *Pattern recognition letters*, 27(8):861–874, 2006.
- Flamary, Rémi, Jrad, Nisrine, Phlypo, Ronald, Congedo, Marco, y Rakotomamonjy, Alain. Mixed-norm regularization for brain decoding. *Computational and mathematical methods in medicine*, 2014, 2014.
- Friedman, Jerome, Hastie, Trevor, y Tibshirani, Rob. Regularization paths for generalized linear models via coordinate descent. *Journal of statistical software*, 33(1):1, 2010.
- Frolov, Alexander A, Húsek, Dušan, Biryukova, Elena V, Bobrov, Pavel Dmitrievitch, Mokienko, Olesya A, y Alexandrov, AV. Principles of motor recovery in post-stroke patients using hand exoskeleton controlled by the brain-computer interface based on motor imagery. *Neural Network World*, 27(1):107, 2017.

- Gersch, Will, Martinelli, F, Yonemoto, J, Low, MD, y Mc Ewan, JA. Automatic classification of electroencephalograms: Kullback-Leibler nearest neighbor rules. *Science*, 205(4402):193–195, 1979.
- Grosse, Ivo, Bernaola-Galván, Pedro, Carpena, Pedro, Román-Roldán, Ramón, Oliver, Jose, y Stanley, H Eugene. Analysis of symbolic sequences using the Jensen-Shannon divergence. *Physical Review E*, 65(4):041905, 2002.
- Gupta, Anjum, Parameswaran, Shibin, y Lee, Cheng-Han. Classification of electroencephalography (EEG) signals for different mental activities using Kullback Leibler (KL) divergence. En *Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on*, págs. 1697–1700. IEEE, 2009.
- Hansen, Per Christian. Analysis of discrete ill-posed problems by means of the L-curve. *SIAM review*, 34(4):561–580, 1992.
- Hastie, Trevor, Tibshirani, Robert, y Buja, Andreas. Flexible discriminant analysis by optimal scoring. *Journal of the American statistical association*, 89(428):1255–1270, 1994.
- Hastie, Trevor, Tibshirani, Robert, y Friedman, Jerome. *The elements of statistical learning: data mining, inference and prediction*. Springer Series in Statistics, 2009.
- Haufe, Stefan, Meinecke, Frank, Görgen, Kai, Dähne, Sven, Haynes, John-Dylan, Blankertz, Benjamin, y Bießmann, Felix. On the interpretation of weight vectors of linear models in multivariate neuroimaging. *Neuroimage*, 87:96–110, 2014.
- Hesse, S, Werner, C, Pohl, M, Rueckriem, S, Mehrholz, Jan, y Lingnau, ML. Computerized arm training improves the motor control of the severely affected arm after stroke: a single-blinded randomized trial in two centers. *Stroke*, 36(9):1960–1966, 2005.
- Hillyard, Steven A. y Kutas, Marta. Electrophysiology of cognitive processing. *Annual review of psychology*, 34(1):33–61, 1983.
- Hodges, Paul W y Bui, Bang H. A comparison of computer-based methods for the determination of onset of muscle contraction using electromyography. *Electroencephalography and Clinical Neurophysiology/Electromyography and Motor Control*, 101(6):511–519, 1996.

- Huang, Ke y Aviyente, Selin. Sparse representation for signal classification. En *Advances in neural information processing systems*, págs. 609–616. 2007.
- Ito, Kazufumi, Jin, Bangti, y Takeuchi, Tomoya. A regularization parameter for nonsmooth tikhonov regularization. *SIAM Journal on Scientific Computing*, 33(3):1415–1438, 2011.
- Jasper, Herbert. Report of the committee on methods of clinical examination in electroencephalography. *Electroencephalogr Clin Neurophysiol*, 10:370–375, 1958.
- Latarjet, Michel y Liard, Alfredo Ruiz. *Anatomía humana*, tomo 1. Ed. Médica Panamericana, 2004.
- Ledesma-Ramirez, Claudia, Bojorges-Valdez, Erik, Yáñez-Suarez, Oscar, Saavedra, Carolina, Bougrain, Laurent, y Gentiletti, Gerardo Gabriel. An open-access P300 speller database. En *Fourth International Brain-Computer Interface Meeting*. 2010.
- Leeb, Robert, Lee, Felix, Keinrath, Claudia, Scherer, Reinhold, Bischof, Horst, y Pfurtscheller, Gert. Brain-computer communication: motivation, aim, and impact of exploring a virtual apartment. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 15(4):473–482, 2007.
- Li, Kun, Raju, Vanitha Narayan, Sankar, Ravi, Arbel, Yael, y Donchin, Emanuel. Advances and challenges in signal analysis for single trial P300-BCI. En *Foundations of Augmented Cognition. Directing the Future of Adaptive Systems*, tomo 2, págs. 87–94. ll, 2011a.
- Li, Kun, Raju, Vanitha Narayan, Sankar, Ravi, Arbel, Yael, y Donchin, Emanuel. Advances and challenges in signal analysis for single trial p300-bci. En *International Conference on Foundations of Augmented Cognition*, págs. 87–94. Springer, 2011b.
- Li, Qing, Lin, Nan, et al. The bayesian elastic net. *Bayesian analysis*, 5(1):151–170, 2010.
- Li, Yuanqing, Yu, Zhu Liang, Bi, Ning, Xu, Yong, Gu, Zhenghui, y Amari, Shun-ichi. Sparse representation for brain signal processing: a tutorial on methods and applications. *IEEE Signal Processing Magazine*, 31(3):96–106, 2014.
- Lin, Jianhua. Divergence measures based on the shannon entropy. *IEEE Transactions on Information Theory*, 37(1):145–151, 1991.

- Lotte, Fabien. Signal processing approaches to minimize or suppress calibration time in oscillatory activity-based brain-computer interfaces. *Proceedings of the IEEE, Institute of Electrical and Electronics Engineers*, 103(6):871–890, 2015.
- Lotte, Fabien, Bougrain, Laurent, Cichocki, Andrzej, Clerc, Maureen, Congedo, Marco, Rakotomamonjy, Alain, y Yger, Florian. A review of classification algorithms for EEG-based Brain-Computer Interfaces: A 10-year update. *Journal of neural engineering*, 2018.
- Lotte, Fabien, Congedo, Marco, Lécuyer, Anatole, Lamarche, Fabrice, y Arnaldi, Bruno. A review of classification algorithms for EEG-based Brain-Computer Interfaces. *Journal of Neural Engineering*, 4(2):R1, 2007.
- Mak, JN, Arbel, Y, Minett, JW, McCane, LM, Yuksel, B, Ryan, D, Thompson, D, Bianchi, L, y Erdogmus, D. Optimizing the P300-based brain-computer interface: current status, limitations and future directions. *Journal of neural engineering*, 8(2):025003, 2011.
- Malouin, Francine, Richards, Carol L, Jackson, Philip L, Laffleur, Martin F, Durand, Anne, y Doyon, Julien. The kinesthetic and visual imagery questionnaire (KVIQ) for assessing motor imagery in persons with physical disabilities: a reliability and construct validity study. *Journal of Neurologic Physical Therapy*, 31(1):20–29, 2007.
- Marchesotti, Silvia, Martuzzi, Roberto, Schurger, Aaron, Blefari, Maria Laura, del Millán, José R, Bleuler, Hannes, y Blanke, Olaf. Cortical and subcortical mechanisms of brain-machine interfaces. *Human brain mapping*, 38(6):2971–2989, 2017.
- Miao, Minmin, Zeng, Hong, Wang, Aimin, Zhao, Changsen, y Liu, Feixiang. Discriminative spatial-frequency-temporal feature extraction and classification of motor imagery EEG: An sparse regression and weighted naïve bayesian classifier-based approach. *Journal of neuroscience methods*, 278:13–24, 2017.
- Mika, Sebastian. *Kernel Fisher Discriminants*. Tesis Doctoral, University of Technology, Berlin, 2002.
- Millán, José del R, Rupp, Rüdiger, Mueller-Putz, Gernot, Murray-Smith, Roderick, Giugliemma, Claudio, Tangermann, Michael, Vidaurre, Carmen, Cincotti, Febo, Kubler, Andrea, Leeb, Robert, et al. Combining brain-computer interfaces and assistive technologies: state-of-the-art and challenges. *Frontiers in neuroscience*, 4:161, 2010.

- Moreno, Pedro J, Ho, Purdy P, y Vasconcelos, Nuno. A Kullback-Leibler divergence based kernel for SVM classification in multimedia applications. En *Advances in neural information processing systems*, pág. None. 2003.
- Mouret, Geoffroy, Brault, Jean-Jules, y Nia, Vahid Partovi. *Generalized elastic net regression*. Groupe d'études et de recherche en analyse des décisions, 2013.
- Müller-Putz, Gernot, Scherer, Reinhold, Brunner, Clemens, Leeb, Robert, y Pfurtscheller, Gert. Better than random: a closer look on BCI results. *International Journal of Bioelectromagnetism*, 10(1):52–55, 2008.
- Novi, Quadrianto, Guan, Cuntai, Dat, Tran Huy, y Xue, Ping. Sub-band common spatial pattern (sbcs) for brain-computer interface. En *Neural Engineering, 2007. CNE'07. 3rd International IEEE/EMBS Conference on*, págs. 204–207. IEEE, 2007.
- Penfield, Wilder y Boldrey, Edwin. Somatic motor and sensory representation in the cerebral cortex of man as studied by electrical stimulation. *Brain*, 60(4):389–443, 1937.
- Peterson, Victoria, Rufiner, Hugo L, y Spies, Rubén D. Ad-hoc gaussian dictionaries for sparse representation of evoked related potentials. En *1.st Pan-american Congress on Computational Mechanics*. 2015a.
- Peterson, Victoria, Rufiner, Hugo L, y Spies, Rubén D. L1-norm regularization for sparse representation and p300 wave detection in brain-computer interfaces. En *V Congreso de Matemática Aplicada, Computacional e Industrial*. 2015b.
- Peterson, Victoria, Rufiner, Hugo L., y Spies, Rubén D. Generalized sparse discriminant analysis for event-related potential classification. *Biomedical Signal Proc. and Control*, 35:70–78, 2017.
- Pfurtscheller, Gert. Functional brain imaging based on ERD/ERS. *Vision research*, 41(10-11):1257–1260, 2001.
- Pfurtscheller, Gert y Aranibar, A. Event-related cortical desynchronization detected by power measurements of scalp EEG. *Clinical Neurophysiology*, 42(6):817–826, 1977.
- Pfurtscheller, Gert y Da Silva, FH Lopes. Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clinical neurophysiology*, 110(11):1842–1857, 1999.
- Pfurtscheller, Gert y Neuper, Christa. Motor imagery activates primary sensorimotor area in humans. *Neuroscience letters*, 239(2-3):65–68, 1997.



- Pfurtscheller, Gert y Neuper, Christa. Motor imagery and direct brain-computer communication. *Proceedings of the IEEE*, 89(7):1123–1134, 2001.
- Pires, Gabriel, Nunes, Urbano, y Castelo-Branco, Miguel. Statistical spatial filtering for a P300-based BCI: tests in able-bodied, and patients with cerebral palsy and amyotrophic lateral sclerosis. *Journal of neuroscience methods*, 195(2):270–281, 2011.
- Ramoser, Herbert, Muller-Gerking, Johannes, y Pfurtscheller, Gert. Optimal spatial filtering of single trial EEG during imagined hand movement. *IEEE transactions on rehabilitation engineering*, 8(4):441–446, 2000.
- Rao, Rajesh P N. *Brain-computer interfacing: an introduction*. Cambridge University Press, 2013.
- Schalk, Gerwin y Mellinger, Jürgen. *A practical guide to brain-computer interfacing with BCI2000: General-purpose software for brain-computer interface research, data acquisition, stimulus presentation, and brain monitoring*. Springer Science & Business Media, 2010.
- Shannon, Claude E. A mathematical theory of communication. *ACM SIGMOBILE Mobile Computing and Communications Review*, 5(1):3–55, 2001.
- Shenoy, Pradeep, Krauledat, Matthias, Blankertz, Benjamin, Rao, Rajesh PN, y Müller, Klaus-Robert. Towards adaptive classification for BCI. *Journal of neural engineering*, 3(1):R13, 2006.
- Singer, Wolf. Synchronization of cortical activity and its putative role in information processing and learning. *Annual review of physiology*, 55(1):349–374, 1993.
- Sjöstrand, Karl, Clemmensen, Line Harder, Larsen, Rasmus, y Ersbøll, Bjarne. SpaSM: A matlab toolbox for sparse statistical modeling. *Journal of Statistical Software Accepted for publication*, 2012.
- Smith, Raymond Carl. *Electroencephalograph based brain computer interfaces*. Tesis Doctoral, Citeseer, 2004.
- Subramanian, S. K, Massie, C. L., Malcolm, M. P., y Levin, M. F. Does provision of extrinsic feedback result in improved motor learning in the upper limb poststroke?. a systematic review of the evidence. *Neurorehabilitation and neural repair*, 24(2):113–124, 2010.

- Thomas, Kavitha P, Guan, Cuntai, Lau, Chiew Tong, Vinod, A Prasad, y Ang, Kai Keng. A new discriminative common spatial pattern method for motor imagery brain-computer interfaces. *IEEE transactions on biomedical engineering*, 56(11):2730–2733, 2009.
- Tibshirani, Robert. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, págs. 267–288, 1996.
- Tibshirani, Ryan Joseph y Taylor, Jonathan E. The solution path of the generalized lasso. *The Annals of Statistics*, 39:1335–1371, 2011.
- Tomioka, Ryota y Müller, Klaus-Robert. A regularized discriminative framework for EEG analysis with application to brain-computer interface. *NeuroImage*, 49(1):415–432, 2010.
- Wolf, Steven L, Winstein, Carolee J, Miller, J Philip, Taub, Edward, Uswatte, Gitendra, Morris, David, Giuliani, Carol, Light, Kathye E, Nichols-Larsen, Deborah, Investigators, Excite, et al. Effect of constraint-induced movement therapy on upper extremity function 3 to 9 months after stroke: the excite randomized clinical trial. *Jama*, 296(17):2095–2104, 2006.
- Wolpaw, Jonathan y Wolpaw, Elizabeth Winter. *Brain-Computer Interfaces: principles and practice*. Oxford University Press, USA, 2012.
- Wolpaw, Jonathan R., Birbaumer, Niels, McFarland, Dennis J., Pfurtscheller, Gert, y Vaughan, Theresa M. Brain-computer interfaces for communication and control. *Clinical neurophysiology*, 113(6):767–791, 2002.
- Wolpaw, Jonathan R y McFarland, Dennis J. Multichannel EEG-based brain-computer communication. *Electroencephalography and clinical Neurophysiology*, 90(6):444–449, 1994.
- Ye, Jieping. Least squares linear discriminant analysis. En *Proceedings of the 24th international conference on Machine learning*, págs. 1087–1093. ACM, 2007.
- Yu, Tianyou, Yu, Zhuliang, Gu, Zhenghui, y Li, Yuanqing. Grouped automatic relevance determination and its application in channel selection for P300 BCIs. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 23(6):1068–1077, 2015.

- Yuan, Ming y Lin, Yi. Model selection and estimation in regression with grouped variables. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 68(1):49–67, 2006.
- Zhang, Rui, Xu, Peng, Chen, Rui, Li, Fali, Guo, Lanjin, Li, Peiyang, Zhang, Tao, y Yao, Dezhong. Predicting inter-session performance of SMR-based brain–computer interface using the spectral entropy of resting-state EEG. *Brain topography*, 28(5):680–690, 2015a.
- Zhang, Yu, Zhou, Guoxu, Jin, Jing, Wang, Xingyu, y Cichocki, Andrzej. Optimizing spatial patterns with sparse filter bands for motor-imagery based brain-computer interface. *Journal of neurosc. methods*, 255:85–91, 2015b.
- Zhang, Yu, Zhou, Guoxu, Jin, Jing, Zhao, Qibin, Wang, Xingyu, y Cichocki, Andrzej. Aggregation of sparse linear discriminant analyses for event-related potential classification in brain-computer interface. *International journal of neural systems*, 24(01):1450003, 2014.
- Zhang, Yu, Zhou, Guoxu, Zhao, Qibin, Jin, Jing, Wang, Xingyu, y Cichocki, Andrzej. Spatial-temporal discriminant analysis for ERP-based brain-computer interface. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 21(2):233–243, 2013.
- Zou, Hui y Hastie, Trevor. Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67(2):301–320, 2005.

**Doctorado en Ingeniería**  
**mención inteligencia computacional, señales y sistemas**

Título de la obra:

**Decodificación de la actividad cerebral mediante  
regularización con penalizantes mixtos**

Autor: Victoria Peterson

Lugar: Santa Fe, Argentina

Palabras Claves:

Interfaces Cerebro-Computadora,  
Análisis Discriminante Regularizado,  
Información Discriminativa a-priori,  
Penalización Mixta,  
Imaginería Motora,  
Potenciales Evocados.