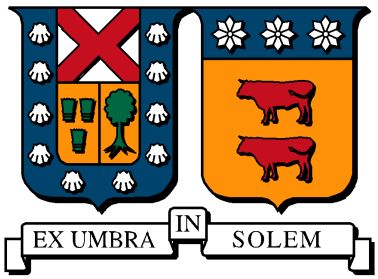
UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA
DEPARTAMENTO DE INGENIERÍA ELECTRÓNICA



Desarrollo de un Pipeline para Remover Artefactos de Masticación en Señales de EEG

Tesis de grado presentada por
Daniela Belén Garrido González

Como requisito parcial para optar al grado de
Magíster en Ciencias de la Ingeniería Electrónica

Director de Tesis
Alejandro Weinstein, Ph.D.

Valparaíso, octubre de 2026

CONSTANCIA DE VALIDACIÓN Y CONFIDENCIALIDAD DE MONOGRAFÍA A REPOSITORIO ACADÉMICO

1.- IDENTIFICACIÓN DEL TRABAJO ACADÉMICO

Tipo de monografía (marcar una opción): ☐ Memoria o trabajo de título ☒ Tesis de Postgrado

Título del trabajo: Desarrollo de un pipeline para remover Artefactos de Masticación en Señales de EEG

Nombre del candidato(a): Daniela Belén Garrido González

Carrera / Grado: Magister en Ciencias de la Ingeniería Electrónica

Campus: Casa Central **Departamento:** Electrónica

2.- VALIDACIÓN DEL PROFESOR GUÍA/DIRECTOR DE TESIS

Yo, Alejandro Weinstein, en mi calidad de profesor(a) guía/director(a) del trabajo académico mencionado anteriormente **DEJO CONSTANCIA** que:

- He revisado esta versión del documento y corresponde a la versión final aprobada del trabajo.
- El trabajo cumple con los requisitos académicos y de formato establecidos por la institución.

3.- EVALUACIÓN DE CONFIDENCIALIDAD POR PROPIEDAD INDUSTRIAL (marcar una opción)

☒ El trabajo **NO contiene** información que amerite confidencialidad y puede ser publicado de inmediato en repositorio con acceso abierto.

☐ El trabajo **CONTIENE** información con potenciales implicancias de propiedad industrial o intelectual y requiere un periodo de confidencialidad (**embargo**) por **(marcar una opción)**:

☐ 6 meses ☐ 12 meses ☐ 2 años ☐ 3 años ☐ 5 años ☐ 10 años

Fundamentación de la necesidad de confidencialidad (obligatorio si se solicita embargo):

4.- FIRMAS

Profesor(a) guía o director(a) de memoria o tesis:

Fecha: 30-10-2026 **Firma:** A. Weinstein D.

Estudiante o Candidato(a):

Fecha: 30 - 10 - 2026 **Firma:**

Este formulario debe ser insertado como página 2 de la memoria o tesis, completado y firmado por estudiante y profesor(a) antes de la entrega en portal PRISMA de Biblioteca USM.

Tesis:
Desarrollo de un Pipeline para Remover Artefactos de Masticación en Señales de EEG

Autor:
Daniela Belén Garrido González

TRABAJO DE TESIS, presentado en cumplimiento parcial de los requisitos para el grado de Magíster en Ciencias de la Ingeniería Electrónica de la Universidad Técnica Federico Santa María, Valparaíso, Chile.

Director de Tesis,
Alejandro Weinstein, Ph.D. _____

Examinador interno,
Matías Zañartu, Ph.D. _____

Examinador Externo,
Alejandro Veloz, Ph.D. _____
(Universidad de Valparaíso)

Agradecimientos

Quiero agradecer, en primer lugar, a mi familia, especialmente a mi hijo Chiumi, mi gatito naranja y mi pequeño rayo de luz. Él me acompañó durante cada uno de estos años en la universidad, estando a mi lado mientras estudiaba, dándome su amor y llenando de luz incluso los días más difíciles. Siempre estaré agradecida con quien dedicó su vida a mi bienestar, acompañándome con un amor tan puro e incondicional.
A mis papás, gracias por apoyarme siempre, por darme ánimos cuando más lo necesité y por recordarme que podía seguir adelante incluso cuando el camino parecía difícil. A mis hermanas y mi Morgana, gracias por ser mi fuerza y la razón por la que sigo aquí.
Agradezco también a cada persona que ha sido parte de mi vida durante este largo camino. A Carlos, por acompañarme y apoyarme durante gran parte de la carrera, por creer en mí y ayudarme a confiar en mis propias capacidades. A mis amigos, a quienes quiero profundamente, gracias por estar presentes, por escucharme, por acompañarme en los buenos y malos momentos y por hacer que este camino fuera más llevadero.
Agradezco también a las amistades que nacieron en las ayudantías, en los espacios de innovación y robótica, en la JIM/BIS, ACATEC y CIAC. Cada una de estas experiencias y las personas que conocí en ellas dejaron una huella en mi vida y contribuyeron a que mi paso por la universidad fuera mucho más que una formación académica. Asimismo, agradezco sinceramente a todos los funcionarios y funcionarias que me acompañaron, orientaron y apoyaron durante este período.
Finalmente, quiero expresar mi especial agradecimiento al profesor Alejandro Weinstein, por guiarme, apoyarme y acompañarme en cada etapa del desarrollo de este trabajo. Gracias por su orientación y por contribuir a mi crecimiento académico y profesional.
A todas las personas que han sido parte de este camino, gracias por acompañarme, por creer en mí y por dejar algo de ustedes en esta etapa de mi vida. Llegar hasta aquí no habría sido lo mismo sin ustedes.
Y a quienes alguna vez han sentido miedo de intentar cumplir sus sueños, quisiera recordarles algo que también me ha acompañado durante este camino:
"No está mal tener miedo. Mis miedos me han traído hasta aquí".
Porque, a veces, ser valiente no significa no tener miedo, sino seguir adelante a pesar de él.

Contenido

Resumen iv

Símbolos y acrónimos v

Lista de figuras viii

Lista de tablas ix

1 Introducción **1**
1.1 Introducción . . . 1
1.2 Motivación . . . 3
1.2.1 Hipótesis . . . 4
1.2.2 Objetivo general . . . 4
1.2.3 Objetivos específicos . . . 4
1.3 Características generales del EEG . . . 5
1.3.1 Sistema internacional 10-20 y su extensión . . . 5
1.4 Artefactos musculares en EEG . . . 6

2 Estado del arte **9**
2.1 Panorama de los métodos de limpieza . . . 9
2.2 Separación ciega de fuentes . . . 10
2.3 Reconstrucción de subespacios de artefactos (ASR) . . . 10
2.4 Filtrado adaptativo con referencias . . . 11
2.5 iCanClean . . . 11
2.6 EEG durante la masticación . . . 12
2.7 Síntesis y enfoque de esta tesis . . . 13

3 Marco teórico **14**
3.1 Notación, preprocesamiento común y principios de comparación . . . 14
3.1.1 Convención y modelo de observación . . . 14
3.1.2 Preprocesamiento común . . . 16
3.1.3 Representaciones de las referencias . . . 16
3.1.4 Principios de comparación . . . 17
3.2 Reconstrucción de subespacios de artefactos (ASR) . . . 17
3.2.1 Supuesto y objetivo . . . 17
3.2.2 Representación matricial . . . 17
3.2.3 Parámetros e interpretación . . . 19

3.2.4 Limitaciones . . . 20
3.3 Filtrado adaptativo (AF) . . . 20
3.3.1 Modelo de señal . . . 20
3.3.2 Actualización normalizada por bloques . . . 21
3.3.3 Limitaciones . . . 22
3.4 Análisis de correlación canónica (CCA) . . . 23
3.4.1 Selección y sustracción . . . 24
3.4.2 Limitaciones . . . 25
3.5 iCanClean . . . 26
3.5.1 Definición . . . 26
3.5.2 Selección, proyección y sustracción . . . 26
3.5.3 Limitaciones . . . 27
3.6 Variante extendida propia . . . 28
3.6.1 Definición . . . 28
3.6.2 Estimación regularizada . . . 28
3.6.3 Limitaciones . . . 29
3.7 Relación entre los métodos y criterios de comparación . . . 29

4 Marco metodológico **31**
4.1 Descripción del conjunto de datos . . . 31
4.1.1 Adquisición . . . 31
4.1.2 Estructura general del conjunto de datos . . . 32
4.1.3 Distribución por experimento . . . 32
4.1.4 Tipos de canales . . . 33
4.1.5 Intervalos de reposo y masticación . . . 34
4.2 Preprocesamiento de señales EEG . . . 34
4.3 Segmentos de calibración y evaluación . . . 37
4.4 Implementación y decisiones por método . . . 39
4.4.1 ASR . . . 39
4.4.2 AF por bloques . . . 41
4.4.3 CCA con rezagos . . . 41
4.4.4 iCanClean . . . 41
4.4.5 Variante extendida propia . . . 42
4.5 Cálculo de métricas, selección y criterios de evaluación . . . 42
4.5.1 Similitud en reposo . . . 43
4.5.2 Reducción de RMS durante masticación . . . 43
4.5.3 Señal eliminada . . . 43
4.5.4 Potencia por banda . . . 44
4.5.5 Interpretación de las métricas . . . 44
4.5.6 Puntaje balanceado . . . 44
4.5.7 Selección de configuraciones y controles . . . 45
4.5.8 Criterios para examinar la hipótesis . . . 47
4.5.9 Agregación y estadística descriptiva . . . 49
4.5.10 Análisis de sensibilidad . . . 49
4.6 Visualización: inspección visual . . . 50
4.7 Prueba semisintética . . . 50

5 Resultados **52**
5.1 Ventanas de evaluación . . . 52
5.2 Resultados de ASR . . . 53
5.2.1 Selección de la configuración . . . 53
5.2.2 Desempeño cuantitativo . . . 53
5.2.3 Inspección visual en AF3 y AF4 . . . 55
5.3 Resultados del AF por bloques . . . 59
5.3.1 Selección de la configuración . . . 59
5.3.2 Desempeño cuantitativo . . . 59
5.3.3 Inspección visual en AF3 y AF4 . . . 62
5.4 Resultados de la CCA con rezagos . . . 64
5.4.1 Selección de la configuración . . . 64
5.4.2 Desempeño cuantitativo . . . 64
5.4.3 Inspección visual en AF3 y AF4 . . . 67
5.5 Resultados de iCanClean . . . 69
5.5.1 Selección de la configuración . . . 69
5.5.2 Desempeño cuantitativo . . . 69
5.5.3 Inspección visual en AF3 y AF4 . . . 73
5.6 Comparación de métodos . . . 76
5.6.1 Comparación global . . . 76
5.6.2 Análisis por banda . . . 76
5.6.3 Controles de especificidad y de actividad ocular . . . 78
5.6.4 Criterios para examinar la hipótesis . . . 79
5.7 Evaluación semisintética . . . 82
5.8 Discusión y limitaciones . . . 83
5.8.1 Compromiso entre la atenuación en masticación y la modificación del reposo 83
5.8.2 Lectura de la etiqueta de la hipótesis . . . 84
5.8.3 Coherencia entre las métricas, la prueba semisintética y la inspección visual . 84
5.8.4 Relación con la literatura . . . 85
5.8.5 Limitaciones . . . 85

6 Conclusiones **88**
6.1 Respuesta a los objetivos . . . 88
6.2 Evaluación de la hipótesis y alcance . . . 89
6.3 Recomendaciones y trabajo futuro . . . 90

Apéndices **91**
A Detalles de implementación y mediciones en los datos de desarrollo . . . 91
B Configuraciones seleccionadas por participante . . . 94
C Resultados complementarios . . . 98
D Análisis de sensibilidad . . . 102
E Trazabilidad . . . 109
F Resultados de la variante extendida . . . 110
G Búsqueda de masticación con el EMG en registros no evaluados . . . 112
H Declaración de uso de herramientas de inteligencia artificial . . . 113

Bibliografía **114**

Resumen

Los artefactos electromiográficos (EMG) de masticación son difíciles de separar del electroencefalograma (EEG), porque su amplio espectro se solapa con bandas de interés neurofisiológico.

Esta tesis evalúa iCanClean, que sustrae del EEG componentes de un análisis de correlación canónica (CCA) cuya correlación al cuadrado con referencias EMG alcanza un umbral, y lo compara con la reconstrucción de subespacios de artefactos (ASR), el filtrado adaptativo (AF) por bloques y la CCA con rezagos.

La hipótesis plantea que iCanClean atenúa más los artefactos de masticación que ASR, AF y CCA, preservando en mayor medida la dinámica espectral y la morfología de la señal en reposo. Con cinco participantes, se examina con criterios descriptivos, que deben cumplirse en al menos cuatro, sobre el compromiso entre reducir la amplitud en masticación y modificar poco el reposo.

Con configuraciones elegidas sin el participante evaluado, ASR obtuvo la mayor reducción media de RMS en masticación (60,6 %), pero también la mayor modificación del reposo (R^2 medio de 61,1 %). El AF, la CCA e iCanClean redujeron la RMS en 14,1 %, 10,6 % y 10,7 %, sobre todo en 1–8 Hz, con R^2 medios en reposo de al menos 97,7 %; el de iCanClean refleja que en varios participantes no actuó en reposo. La CCA e iCanClean dependieron de sus propias referencias para atenuar. Con el EOG vertical, lo que ambos y el AF retiraron en masticación compartió menos varianza en 1–8 Hz que la entrada; con el horizontal, más en iCanClean y, salvo en E1M2_R2, en la CCA. Sin embargo, el criterio de atenuación no se cumplió frente a ningún comparador, y la hipótesis resultó «no respaldada», etiqueta que no califica la limpieza. Una variante propia de regresión ridge no superó el puntaje balanceado medio de iCanClean. Estas métricas describen cambios del EEG observado, no la preservación de la actividad cerebral.

Palabras clave: electroencefalografía, masticación, artefactos electromiográficos, limpieza de señales, iCanClean, filtrado adaptativo, ASR, CCA.

Símbolos y acrónimos

Los símbolos y acrónimos que se indican a continuación se utilizan en la presente tesis. La Sección 3.1 explica la convención de notación, y los símbolos que requieren una definición más rigurosa se definen en los Capítulos 3 y 4, justo antes de utilizarse. Los símbolos que intervienen solo en una deducción se definen en ese lugar y no se incluyen en la lista.

Acrónimos

Acrónimo	Significado
AC3E	Centro Avanzado de Ingeniería Eléctrica y Electrónica (*Advanced Center for Electrical and Electronic Engineering*)
AF	Filtrado adaptativo (*Adaptive Filtering*)
ASR	Reconstrucción de subespacios de artefactos (*Artifact Subspace Reconstruction*)
BSS	Separación ciega de fuentes (*Blind Source Separation*)
CCA	Análisis de correlación canónica (*Canonical Correlation Analysis*)
CMS	Electrodo de modo común (*Common Mode Sense*)
DE	Desviación estándar
DRL	Electrodo de pierna derecha impulsada (*Driven Right Leg*)
EEG	Electroencefalografía (la técnica) o electroencefalograma (la señal registrada), según el contexto
EEMD	Descomposición empírica en modos por conjuntos (*Ensemble Empirical Mode Decomposition*)
EMG	Electromiografía o electromiograma, según el contexto
EOG	Electrooculografía o electrooculograma, según el contexto
FASTER	Umbralización estadística completamente automatizada para el rechazo de artefactos del EEG (*Fully Automated Statistical Thresholding for EEG artifact Rejection*)
FIR	Respuesta finita al impulso (*Finite Impulse Response*)
ICA	Análisis de componentes independientes (*Independent Component Analysis*)
ICLabel	Clasificador automático de componentes independientes del EEG
iCanClean	Algoritmo publicado de limpieza basado en la CCA entre el EEG y referencias de ruido (Sección 3.5)
LMS	Mínimos cuadrados medios (*Least Mean Squares*)
MEMD	Descomposición empírica en modos multivariada (*Multivariate Empirical Mode Decomposition*)
NIPALS	Mínimos cuadrados parciales iterativos no lineales (*Nonlinear Iterative Partial Least Squares*)
NMSE	Error cuadrático medio normalizado (*Normalized Mean Square Error*)
PCA	Análisis de componentes principales (*Principal Component Analysis*)
PREP	Secuencia estandarizada de preprocesamiento del EEG (*PREP pipeline*)
PSD	Densidad espectral de potencia (*Power Spectral Density*)
RLS	Mínimos cuadrados recursivos (*Recursive Least Squares*)
RMS	Raíz cuadrática media (*Root Mean Square*)
RRMSE	Raíz del error cuadrático medio relativo (*Relative Root Mean Square Error*)
SHA	Algoritmo de resumen seguro (*Secure Hash Algorithm*)
SNR	Relación señal a ruido (*Signal-to-Noise Ratio*)
SSA	Análisis espectral singular (*Singular Spectrum Analysis*)
wICA	ICA mejorado con *wavelets* (*wavelet-enhanced ICA*)

Símbolos

Las matrices y los vectores se escriben en negrita, y los escalares, en cursiva. En todas las matrices de señales, las muestras se ordenan por filas y los canales por columnas. El índice n denota la muestra; c, el canal del EEG, y q, la referencia EMG. La traspuesta se denota por $(\cdot)^\mathsf{T}$; la pseudoinversa, por $(\cdot)^+$, y la norma euclidiana, por $\|\cdot\|$. El subíndice cal indica calibración; MC, centrado por columnas; z, estandarizado por columnas con los estadísticos de la calibración; w, una ventana de ASR, y b, un bloque del filtro adaptativo. En el análisis, las bandas de frecuencia se designan por su nombre y su rango: delta (1–4 Hz), theta (4–8 Hz), alfa (8–13 Hz), beta (13–30 Hz) y gamma (30–45 Hz).

Símbolo	Significado
Señales y modelo	
$\boldsymbol{X}$, $\boldsymbol{x}[n]$, $x_c[n]$	EEG observado, $\boldsymbol{X} \in \mathbb{R}^{T\times C}$; fila n y canal c
T, C	Número de muestras y número de canales
T_{cal}	Número de muestras del segmento de calibración
$\bar{\boldsymbol{x}}_{\mathrm{cal}}$, $\boldsymbol{\sigma}_{\mathrm{cal}}$	Medias y desviaciones estándar por canal del segmento de calibración
$\boldsymbol{S}$, $\boldsymbol{A}$, $\boldsymbol{\varepsilon}$	Actividad neuronal de interés, artefacto y residuo del modelo $\boldsymbol{X} = \boldsymbol{S} + \boldsymbol{A} + \boldsymbol{\varepsilon}$
$\widehat{\boldsymbol{A}}$, $\widehat{a}_c[n]$	Artefacto estimado por un método y su columna para el canal c
$\boldsymbol{X}_{\mathrm{clean}}$	Señal corregida
γ	Factor de escala con que se sustrae el artefacto estimado
$\boldsymbol{R}$, $\boldsymbol{r}[n]$, $r_q[n]$	Referencias EMG, $\boldsymbol{R} \in \mathbb{R}^{T\times Q_{\mathrm{EMG}}}$; fila n y referencia q
Q_{EMG}	Número de referencias EMG: 2 derivaciones bipolares en la configuración principal o 4 en los análisis de sensibilidad con canales monopolares
$\breve{r}_q[n]$	EMG de la referencia q filtrado en 20–120 Hz, con rechazo de banda en 50 y 100 Hz, a partir del cual se calcula la envolvente
$\boldsymbol{R}_{\mathrm{env}}$, $r_q^{\mathrm{env}}[n]$	Envolvente de cada referencia, calculada a partir del EMG de 20–120 Hz, centrada y filtrada luego en 1–45 Hz
f_{env}	Frecuencia de corte del filtro de la envolvente
$\tilde{\boldsymbol{R}}$, $\tilde{Q}$	Referencia ampliada con la envolvente, $\tilde{\boldsymbol{R}} = [\boldsymbol{R}, \boldsymbol{R}_{\mathrm{env}}]$, y su número de columnas; sin envolvente, $\tilde{\boldsymbol{R}} = \boldsymbol{R}$
$\boldsymbol{R}^{(j)}$	Referencias desplazadas j muestras: su fila n es $\boldsymbol{r}[n-j]$, con ceros fuera de rango; $j>0$ corresponde al pasado y $j<0$ al futuro
$\boldsymbol{R}_{\mathrm{lag}}$, L	Referencias con rezagos, $\boldsymbol{R}_{\mathrm{lag}} = [\boldsymbol{R}^{(-L)}, \ldots, \boldsymbol{R}^{(L)}]$, y rezago máximo
$\tilde{\boldsymbol{R}}_{\mathrm{lag}}$	Referencias ampliadas con envolvente y rezagos, $\tilde{\boldsymbol{R}}_{\mathrm{lag}} = [\tilde{\boldsymbol{R}}^{(-L)}, \ldots, \tilde{\boldsymbol{R}}^{(L)}]$; sin envolvente, $\tilde{\boldsymbol{R}}_{\mathrm{lag}} = \boldsymbol{R}_{\mathrm{lag}}$ (variante extendida)
Reconstrucción de subespacios de artefactos (ASR)	
$\boldsymbol{\Sigma}_{\mathrm{cal}}$, $\boldsymbol{\Sigma}_*$	Covarianza muestral de la calibración y su versión robusta (mediana geométrica)
$\boldsymbol{M}$	Raíz de la covarianza robusta, $\boldsymbol{M} = \boldsymbol{\Sigma}_*^{1/2}$
$\boldsymbol{\Sigma}_w$	Matriz local de segundos momentos de la señal ponderada en la ventana w
$\boldsymbol{\Phi}_{\mathrm{cal}}$, $\boldsymbol{\Phi}_w$	Vectores propios de la calibración y de la ventana w
k	Tolerancia de ASR
$\boldsymbol{\Theta}$	Matriz de umbrales
$\delta_{w,j}$, $\boldsymbol{D}_w$	Indicador de conservación de la componente j, $\delta_{w,j} \in \{0,1\}$, y matriz $\boldsymbol{D}_w = \mathrm{diag}(\boldsymbol{\delta}_w)$
f_{ASR}, m	Fracción dimensional y número de direcciones que fija, $m = \mathrm{redondeo}(C\, f_{\mathrm{ASR}})$; en cada actualización pueden reconstruirse hasta $m+1$ direcciones
$\boldsymbol{G}_w$	Matriz de reconstrucción, $\boldsymbol{G}_w = \boldsymbol{M}(\boldsymbol{D}_w\boldsymbol{\Phi}_w^\mathsf{T}\boldsymbol{M})^+\boldsymbol{\Phi}_w^\mathsf{T}$
Filtrado adaptativo (AF) por bloques	
L_{AF}	Memoria del filtro, en muestras
$\tilde{x}_c[n]$, $\tilde{\boldsymbol{X}}_b$	Objetivo normalizado del canal c y su matriz en el bloque b
$\boldsymbol{\psi}[n]$, $\boldsymbol{\Psi}_b$	Vector de regresores y su matriz en el bloque b
$\tilde{a}_c[n]$, $\tilde{\boldsymbol{A}}_b$	Predicción normalizada del artefacto en el canal c y su matriz en el bloque b

Símbolo	Significado
$\boldsymbol{W}_b$, $\boldsymbol{w}_c$	Pesos del filtro vigentes al inicio del bloque b y pesos del canal c
$e_c[n]$	Residual del AF (señal corregida, en unidades del EEG) del canal c, $e_c[n] = x_c[n] - \widehat{a}_c[n]$
$\tilde{e}_c[n]$, $\boldsymbol{E}_b$	Residual normalizado del AF del canal c y su matriz en el bloque b
η_b	Energía de normalización del bloque b
μ	Paso de adaptación
ν	Fuga
B	Avance entre bloques, en muestras
Análisis de correlación canónica (CCA con rezagos e iCanClean)	
$\boldsymbol{\Omega}_X$, $\boldsymbol{\Omega}_R$	Pesos canónicos del EEG y de las referencias
$\boldsymbol{U}$, $\boldsymbol{V}$	Variables canónicas del EEG y de las referencias: $\boldsymbol{U} = \boldsymbol{X}_{\mathrm{MC}}\boldsymbol{\Omega}_X$ y $\boldsymbol{V} = \boldsymbol{R}_{\mathrm{MC}}\boldsymbol{\Omega}_R$ en iCanClean; $\boldsymbol{U} = \boldsymbol{X}_z\boldsymbol{\Omega}_X$ y $\boldsymbol{V} = \boldsymbol{R}_{\mathrm{lag},z}\boldsymbol{\Omega}_R$ en la CCA con rezagos
ρ_i	Correlación canónica del par i; en la CCA con rezagos se recalcula en el segmento que se limpia y puede ser negativa
n_{comp}	Número de componentes de la CCA con rezagos
τ	Umbral sobre $\lvert\rho_i\rvert$ en la CCA con rezagos
θ	Umbral sobre ρ_i^2 en iCanClean
$\mathcal{B}$, $\lvert\mathcal{B}\rvert$	Conjunto de componentes seleccionadas y su número
$\boldsymbol{Z}$	Actividad seleccionada: $\boldsymbol{U}_{:,\mathcal{B}}$ (fuente $\boldsymbol{U}$) o $\boldsymbol{V}_{:,\mathcal{B}}$ (fuente $\boldsymbol{V}$)
$\boldsymbol{P}$	Proyección por mínimos cuadrados de la actividad seleccionada (iCanClean)
$\boldsymbol{P}_\alpha$, α	Proyección ridge de la CCA con rezagos y su regularización
g_c, $g_{\mathrm{máx}}$	Ganancia del canal c, $g_c \in [0, g_{\mathrm{máx}}]$, y su valor máximo
$\Delta_{\mathrm{máx}}$	Aumento relativo de RMS a partir del cual se revierte la corrección de un canal
Variante extendida	
$\boldsymbol{B}_\lambda$	Coeficientes de la regresión ridge con referencias rezagadas
λ	Regularización de la variante extendida
Métricas y evaluación	
$x[n]$, $y[n]$, N	Entrada y salida de un canal en las métricas, y número de muestras de las ventanas de evaluación de una condición
R^2, CC	Similitud, $R^2 = 100\, r_{xy}^2$ (%), y coeficiente de correlación, $\mathrm{CC} = r_{xy}$; con subíndice se indica la condición, como en R^2_{reposo}
Q_{RMS}	Razón RMS, $Q_{\mathrm{RMS}} = \mathrm{RMS}(y)/\mathrm{RMS}(x)$
$\Delta\mathrm{RMS}$	Reducción RMS, $\Delta\mathrm{RMS} = 100(1 - Q_{\mathrm{RMS}})$ (%); durante la masticación, $\Delta\mathrm{RMS}_{\mathrm{mast}}$
E_{rel}	Amplitud relativa eliminada, $E_{\mathrm{rel}} = \lVert x - y\rVert / \lVert x\rVert$
$\widehat{P}_x(f)$, $\widehat{P}_y(f)$	Densidades espectrales de potencia de la entrada y de la salida, estimadas con el método de Welch
$\mathcal{F}$	Banda de frecuencia, $\mathcal{F} = [f_1, f_2]$, con ambos bordes incluidos
$P_{\mathrm{in}}(\mathcal{F})$, $P_{\mathrm{out}}(\mathcal{F})$	Potencia de la entrada y de la salida en la banda $\mathcal{F}$; su razón es $P_{\mathrm{out}}/P_{\mathrm{in}}$
S_{bal}	Puntaje balanceado, $S_{\mathrm{bal}} = \frac{1}{2}R^2_{\mathrm{reposo}} + \frac{1}{2}\Delta\mathrm{RMS}_{\mathrm{mast}}$
$S_{\mathrm{bal,banda}}$	Puntaje de sensibilidad, que reemplaza $\Delta\mathrm{RMS}_{\mathrm{mast}}$ por la reducción de potencia en 20–45 Hz durante la masticación
p, p', $\mathcal{P}$	Participante evaluado, otro participante y conjunto de participantes; $\mathcal{P} \setminus \{p\}$ son los participantes con que se selecciona la configuración de p
ζ, $\zeta^*(p)$	Configuración de un método y configuración elegida para el participante p
$\boldsymbol{X}_{\mathrm{rep}}$, $\boldsymbol{X}_{\mathrm{mast}}$	Pieza de reposo de evaluación que sirve de objetivo conocido (un reposo observado, no una señal cerebral pura) y contaminación de masticación que se le suma en la prueba semisintética
$\mathrm{RRMSE}_{\mathrm{t}}$, $\mathrm{RRMSE}_{\mathrm{f}}$	Error relativo temporal y espectral de la salida respecto de $\boldsymbol{X}_{\mathrm{rep}}$

Lista de figuras

1.1 Ejemplos de artefactos oculares y musculares en un registro de EEG 2
1.2 Posiciones nominales de los 64 canales EEG 6
1.3 Artefactos musculares de masticación en canales frontales 8

3.1 Problema común de la tesis 15
3.2 Flujo de ASR 20
3.3 AF por bloques 23
3.4 CCA con rezagos 25
3.5 Etapas de iCanClean 27
3.6 Variante extendida propia 29

4.1 Disposición de los electrodos EOG y EMG 33
4.2 Secuencia de preprocesamiento común 35
4.3 Reservas de calibración y ventanas de evaluación 38

5.1 ASR: compromiso entre similitud en reposo y reducción de RMS en la selección . . . 54
5.2 ASR: métricas por canal 56
5.3 ASR: inspección visual en AF3 y AF4, trazas y PSD 57
5.4 ASR: espectrogramas en AF3 y AF4 58
5.5 AF por bloques: compromiso entre similitud en reposo y reducción de RMS en la selección 60
5.6 AF por bloques: métricas por canal 61
5.7 AF por bloques: inspección visual en AF3 y AF4, trazas y PSD 62
5.8 AF por bloques: espectrogramas en AF3 y AF4 63
5.9 CCA con rezagos: compromiso entre similitud en reposo y reducción de RMS en la selección 65
5.10 CCA con rezagos: métricas por canal 66
5.11 CCA con rezagos: inspección visual en AF3 y AF4, trazas y PSD 67
5.12 CCA con rezagos: espectrogramas en AF3 y AF4 68
5.13 iCanClean: compromiso entre similitud en reposo y reducción de RMS en la selección 71
5.14 iCanClean: métricas por canal 72
5.15 iCanClean: inspección visual en AF3 y AF4, trazas y PSD 74
5.16 iCanClean: espectrogramas en AF3 y AF4 75
5.17 Razón de potencia por banda, media de los participantes 77

Lista de tablas

2.1 Comparación de los enfoques de limpieza revisados 13

3.1 Resumen de los métodos comparados 30

4.1 Intervalos de masticación y de reposo 34
4.2 Configuraciones exploradas por método 40
4.3 Métricas de evaluación 45

5.1 Ventanas de evaluación y calibración efectiva de ASR 53
5.2 ASR: métricas por participante 55
5.3 AF por bloques: métricas por participante 59
5.4 CCA con rezagos: métricas por participante 64
5.5 iCanClean: ventanas con componentes retiradas 70
5.6 iCanClean: métricas por participante 70
5.7 Comparación global de los métodos 76
5.8 Fracción de la atenuación que se obtiene con referencias de otro participante . . . 78
5.9 Control de actividad ocular en 1–8 Hz 80
5.10 Criterios para examinar la hipótesis 81
5.11 Matriz de los criterios para examinar la hipótesis 81

B.1 Configuración elegida para cada participante 94
B.2 Resolución de la selección para cada participante 95
B.3 Identificadores de las configuraciones de cada método 96
B.4 Configuraciones informadas sin selección y sus métricas medias 97
C.1 iCanClean: discontinuidades en los bordes entre ventanas 99
C.2 Razón de potencia por banda por participante 99
C.3 Controles de especificidad por participante 101
C.4 Evaluación semisintética 102
D.1 ASR: sensibilidad a detalles del procesamiento 103
D.2 Sensibilidad a la referencia EMG 103
D.3 Criterio 1 en las ventanas independientes del diseño 105
D.4 Métricas sin la ventana más próxima a la reserva 106
D.5 Criterios en las sensibilidades de diseño 107
D.6 iCanClean según la duración de la ventana: métricas 107
D.7 iCanClean: ventanas con selección y nivel de azar de ρ_1^2 108
E.1 Identificación de los archivos, del software y de los resultados 109
F.1 Variante extendida: métricas y controles de especificidad 111

CAPÍTULO

1 Introducción

1.1. Introducción

El electroencefalograma (EEG) es una técnica no invasiva ampliamente utilizada para estudiar la dinámica cerebral, debido a su alta resolución temporal, del orden de milisegundos, y a su aplicabilidad en contextos clínicos y de investigación [1, 2]. Sin embargo, una limitación práctica es su susceptibilidad a fuentes de ruido no neuronal, lo que afecta la calidad y la interpretabilidad de los registros.

En condiciones reales de adquisición, las señales EEG a menudo se encuentran contaminadas por artefactos, tales como los producidos por las contracciones musculares al morder, masticar o fruncir el ceño, o por los movimientos al caminar. Los artefactos musculares, en particular, presentan un desafío considerable, ya que su espectro se superpone con el de la actividad cerebral, lo que los hace difíciles de separar de ella [2, 3]. En términos generales, los artefactos corresponden a componentes presentes en el registro que no representan directamente la actividad neuronal de interés. Pueden originarse en fuentes fisiológicas no cerebrales, como los movimientos oculares, la actividad cardíaca o las contracciones musculares, o en factores externos, como el desplazamiento de los electrodos, el movimiento del sujeto o la interferencia eléctrica [2, 4]. Además, los artefactos pueden imitar actividad cognitiva o patológica y, por lo tanto, sesgar la interpretación visual y el diagnóstico en investigaciones clínicas [5]. En la Figura 1.1 se muestran ejemplos de artefactos oculares y musculares en uno de los registros utilizados en esta tesis.

Esta situación impacta directamente el análisis de la señal. Si la limpieza es insuficiente, las métricas de potencia, conectividad o detección de eventos pueden quedar sesgadas. Por otro lado, si la limpieza es excesiva, existe el riesgo de eliminar actividad cerebral relevante. Por ejemplo, en un estudio sobre la localización del inicio de crisis epilépticas en el EEG de superficie, ambos errores

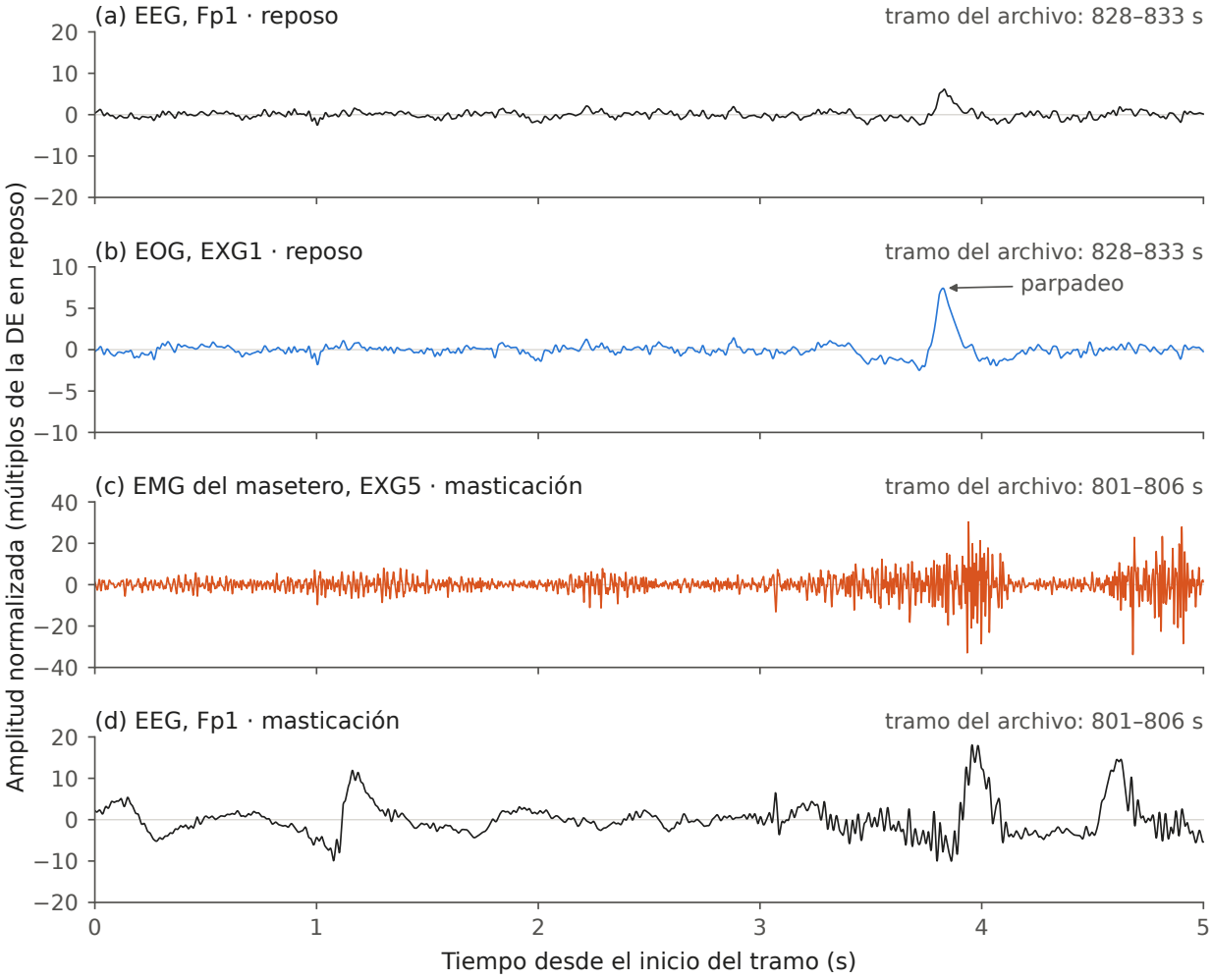


Figura 1.1: Ejemplos de artefactos oculares y musculares en el registro de E1M1_R2. (a) EEG de Fp1 en reposo. (b) Electrooculograma (EOG; EXG1) en el mismo tramo; la flecha señala un probable parpadeo, cerca de 3,8 s, visible también en Fp1. (c) Electromiograma (EMG) del masetero izquierdo (EXG5) durante la masticación. (d) EEG de Fp1 en ese tramo, con actividad de alta frecuencia durante las ráfagas y deflexiones lentas de gran amplitud (Sección 1.4). Reposo: 828–833 s; masticación: 801–806 s. Filtrado en 1–45 Hz (EEG y EOG) y 20–120 Hz (EMG), referencia de modo común (CMS) y cada traza dividida por su desviación estándar (DE) en el reposo. Elaboración propia a partir de registros de Sebastián Espinoza, laboratorio de neurociencia del AC3E (Sección 4.1).

llevaron a lateralizaciones incorrectas en dos pacientes [6]. En este contexto, reducir la actividad muscular no es suficiente. El método también debe evitar modificaciones innecesarias de la señal EEG, especialmente durante los segmentos en que el sujeto se encuentra en reposo.

1.2. Motivación

La masticación constituye una condición experimental de interés debido a la actividad muscular asociada al movimiento mandibular. Las contracciones de los músculos involucrados pueden generar artefactos de amplitud considerable y contenido espectral amplio [7], capaces de contaminar simultáneamente múltiples canales EEG [8]. Además, investigaciones recientes han reportado que esta actividad rítmica se asocia con cambios en las funciones cognitivas y en la dinámica cerebral.

En un estudio difundido como *preprint*, todavía sin revisión por pares, Espinoza et al. reportaron que masticar durante una tarea de memoria de trabajo se asoció con respuestas más rápidas y con un acoplamiento fase-amplitud: la amplitud de las oscilaciones θ (4–8 Hz) frontocentrales se moduló según la fase del ciclo masticatorio [7]. En otro estudio del mismo grupo, ya publicado, masticar durante un minuto antes de tareas de atención y de memoria de trabajo mejoró el desempeño y aumentó la potencia y la coherencia θ frontocentrales y la sincronización entre electrodos frontales y parietales [9]. Los autores interpretan estos resultados como una sincronización de la actividad θ inducida por el ritmo masticatorio [7, 9]. Sin embargo, ese acoplamiento depende de la limpieza del registro, porque la actividad muscular mandibular y facial oscurece especialmente las oscilaciones θ [7], y un artefacto residual sincronizado con la masticación podría producir un patrón similar.

Por otra parte, la masticación se ha relacionado con la salud cerebral. Según una revisión narrativa, el acto masticatorio conduce una gran cantidad de información sensorial al sistema nervioso central y se asocia con un aumento de la oxigenación sanguínea en el hipocampo y en la corteza prefrontal, estructuras fundamentales para el aprendizaje y la memoria [10]. Con evidencia obtenida en buena medida en modelos animales, la misma revisión plantea que una correcta función masticatoria evita el deterioro del hipocampo asociado a la pérdida dental o a la desarmonía oclusal, y presenta la masticación como una posible herramienta para prevenir disfunciones cognitivas [10]. Comprender esta relación requiere registros EEG interpretables durante el propio acto masticatorio, que es precisamente cuando la contaminación muscular es mayor.

El estudio de la actividad electroencefalográfica durante la masticación constituye un desafío, debido a que los artefactos electromiográficos que introduce esta condición pueden superponerse en frecuencia y en tiempo con componentes de interés fisiológico [7]. Esta superposición dificulta la interpretación de la actividad EEG registrada durante la masticación y plantea la necesidad de evaluar métodos capaces de reducir la contaminación muscular sin modificar excesivamente la señal registrada. Tradicionalmente, la ausencia de procedimientos de limpieza robustos para este nivel de contaminación ha dificultado el estudio de la actividad neurocognitiva mientras se mastica [7]. Por lo tanto, el desarrollo y la evaluación de algoritmos de atenuación de artefactos durante la masticación son una necesidad metodológica para avanzar en la neurociencia cognitiva.

Desde el punto de vista del procesamiento de señales, el interés principal de la masticación en este trabajo corresponde a la contaminación electromiográfica de los músculos del movimiento mandibular. Para abordarla, se evalúa el algoritmo iCanClean en su formulación publicada, que utiliza el análisis de correlación canónica (CCA) para sustraer la actividad del EEG más correlacionada con

referencias de ruido [11, 12], aquí señales electromiográficas (EMG) de los maseteros. Este algoritmo se compara, con un procedimiento común, con la reconstrucción de subespacios de artefactos (ASR), el filtrado adaptativo (AF) por bloques y la CCA con rezagos. Además, se desarrolla una variante extendida propia, una regresión ridge sobre las referencias EMG rezagadas, distinta de iCanClean y evaluada por separado, sin intervenir en el examen de la hipótesis.

1.2.1. Hipótesis

La hipótesis de este trabajo es la siguiente: el algoritmo iCanClean logra una mayor atenuación de artefactos miogénicos derivados de la masticación en registros EEG, preservando en mayor medida la dinámica espectral y la morfología de la señal en condición de reposo, en comparación con los métodos ASR, AF y CCA.

En esta tesis, la hipótesis se refiere al algoritmo iCanClean publicado y se examina de forma exploratoria, sin un contraste estadístico confirmatorio, con criterios descriptivos sobre la atenuación en masticación y la modificación del reposo (Sección 4.5.8).

1.2.2. Objetivo general

Evaluar el desempeño de iCanClean para la atenuación de artefactos de masticación en señales EEG y compararlo, mediante un conjunto común de métricas de evaluación, con los métodos ASR, AF y CCA, y examinar con el mismo procedimiento una variante extendida propia.

1.2.3. Objetivos específicos

1. Implementar e integrar el algoritmo iCanClean en un entorno de análisis en Python común a todos los métodos, con el EMG de los maseteros como referencias y según su formulación publicada: análisis de correlación canónica entre el EEG y las referencias, selección de componentes con un umbral sobre la correlación canónica al cuadrado y proyección por mínimos cuadrados de la actividad seleccionada.

2. Desarrollar, como contribución propia, una variante extendida basada en una regresión ridge con referencias EMG rezagadas, con y sin la envolvente de esas referencias, y definir los parámetros asociados a la regularización.

3. Implementar las técnicas de comparación (filtrado adaptativo por bloques, reconstrucción de subespacios de artefactos mediante asrpy y CCA con rezagos) en el mismo entorno de trabajo.

4. Evaluar cuantitativa y cualitativamente el desempeño de los métodos en segmentos de reposo y de masticación, mediante métricas de similitud con la señal de entrada, reducción de amplitud, error relativo y potencia por banda, con la configuración de cada método elegida sin utilizar los datos del participante evaluado, y complementar esta evaluación con una prueba semisintética.

1.3. Características generales del EEG

El EEG registra las variaciones de potencial eléctrico generadas por la actividad neuronal mediante electrodos ubicados sobre el cuero cabelludo. Su amplitud típica es baja, del orden de decenas de microvoltios [4], mientras que la de los artefactos puede ser mucho mayor [2]. Por eso, aunque refleja con gran resolución temporal la dinámica cerebral, el EEG resulta vulnerable a interferencias de origen externo, asociadas al instrumento de medición, y de origen interno, asociadas al propio sujeto [5].

Para caracterizar la actividad EEG se utilizan habitualmente las bandas de frecuencia δ, por debajo de 4 Hz; θ, de 4 a 8 Hz; α, de 8 a 13 Hz; β, de 13 a 30 Hz; y γ, de 30 a 80 Hz [4]. En esta tesis, el análisis se restringe al intervalo de 1 a 45 Hz, por lo que las bandas δ y γ se consideran solo dentro de ese intervalo (Sección 4.5).

En el EEG de cuero cabelludo, la actividad de alta frecuencia es más difícil de registrar. La banda gamma se observa con mayor frecuencia mediante electrodos intracraneales [4], y su amplitud en el cuero cabelludo puede ser varios órdenes de magnitud menor que la de la actividad muscular registrada en electrodos temporales [13]. Esta característica obliga a balancear la conservación de información fisiológica con la necesidad de atenuar ruido y artefactos de alta frecuencia.

1.3.1. Sistema internacional 10-20 y su extensión

El sistema internacional 10-20, recomendado por la Federación Internacional de Neurofisiología Clínica, es un estándar para ubicar los electrodos sobre el cuero cabelludo [4]. En este sistema, las posiciones se determinan midiendo la cabeza a partir de cuatro puntos de referencia anatómicos externos y tomando el 10 % o el 20 % de esas distancias, de donde proviene su nombre [4, 14]. En su forma básica, el sistema define 21 posiciones [15]. Para registros de mayor densidad, se extendió a los sistemas 10-10 y 10-5, que agregan posiciones intermedias y permiten definir más de 300 ubicaciones [15, 16]. Los montajes de 64 a 256 electrodos se consideran de alta densidad, y la estandarización de las posiciones resulta esencial en la clínica y en la investigación [14], porque facilita la reproducibilidad de los estudios y la comparación entre registros.

El registro EEG utilizado en este trabajo cuenta con 64 canales, cuyos rótulos siguen la nomenclatura del sistema 10-20 extendido. En la Figura 1.2 se muestran sus posiciones nominales, es decir, las que corresponden a cada rótulo en una plantilla estándar, y no una medición en cada participante (Sección 4.1).

La utilidad del sistema 10-20 extendido no se limita a la colocación de los electrodos. También proporciona una estructura para interpretar la distribución topográfica de la actividad cerebral y de los artefactos. En estudios de limpieza, esta organización espacial resulta importante porque algunos artefactos, en especial los musculares, afectan de manera desigual a las distintas zonas del cuero cabelludo. Por ejemplo, la contaminación producida por los músculos frontal y temporal es mayor en la periferia del cuero cabelludo, cerca de los músculos activos [17].

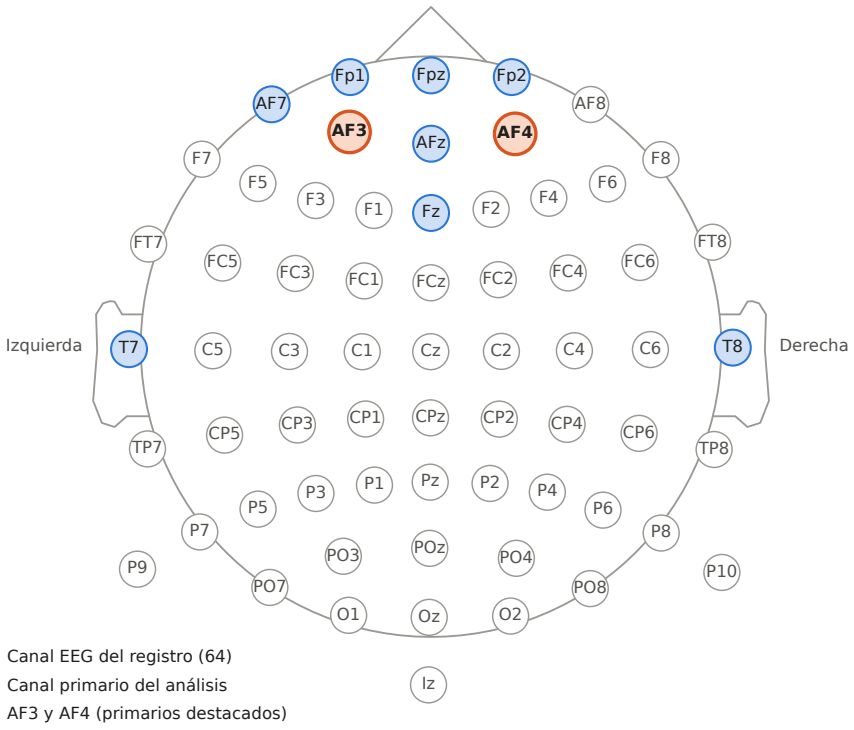


Figura 1.2: Posiciones nominales de los 64 canales EEG del registro, vistas desde arriba y con la nariz hacia el frente, según la plantilla estándar 10-5 de MNE-Python [1] en proyección azimutal equidistante. P9, P10 e Iz quedan fuera del contorno de la cabeza. Los diez canales primarios del análisis (Sección 4.5) están en azul, salvo AF3 y AF4, en naranja, que se usan en la inspección visual (Sección 4.6). Elaboración propia.

En esta tesis, todos los registros comparten los mismos rótulos de canal, lo que permite referirse a cada canal de manera coherente entre participantes y distinguir los canales EEG de los canales auxiliares de referencia muscular.

1.4. Artefactos musculares en EEG

Los artefactos musculares constituyen una fuente relevante de contaminación en registros EEG, especialmente cuando se producen contracciones de músculos cercanos a los electrodos. La actividad eléctrica de los músculos de la cabeza, la cara y el cuello se propaga por conducción de volumen. Por esta razón, la contaminación no se limita a los electrodos adyacentes al músculo, sino que puede extenderse a todo el cuero cabelludo [17, 18].

Una de las principales dificultades de estos artefactos es que presentan características no estacionarias [2]. Su amplitud, su frecuencia dominante y su ancho de banda pueden cambiar según el músculo involucrado, la intensidad y la lateralidad de la contracción y el sexo del participante [13], y su composición espectral también varía con la fatiga [18]. Esto los vuelve más

complejos de tratar que los artefactos de dinámica más regular o predecible, como los movimientos oculares o los parpadeos [18].

Desde el punto de vista espectral, la actividad muscular ocupa una banda amplia, que en promedio se extiende desde frecuencias muy bajas hasta más de 200 Hz [17], y puede superponerse con todas las bandas EEG de interés [18]. La banda γ es especialmente vulnerable: en dos sujetos registrados con y sin bloqueo neuromuscular, la mayor parte del EEG de cuero cabelludo sobre 20 Hz resultó ser de origen muscular [19]. También pueden verse afectadas las bandas β y α, ya que incluso contracciones débiles producen actividad detectable sobre 12 Hz cerca del vértex, y sobre 8 Hz en algunos individuos [17], e incluso las bandas θ y δ, porque se han informado efectos miogénicos desde unos 2 Hz [18]. De este modo, la actividad muscular resulta difícil de distinguir de la actividad cerebral únicamente por criterios espectrales. Este solapamiento limita la efectividad de las estrategias basadas exclusivamente en filtros de frecuencia, que eliminan junto con el artefacto información cerebral de interés [20]. Por lo tanto, resulta necesario considerar métodos que utilicen información espacial, temporal o referencias auxiliares [5].

En contextos experimentales con masticación o tensión facial, la contaminación muscular puede ser suficientemente intensa como para modificar las métricas que se calculan sobre la señal, como la potencia espectral, la conectividad o el análisis de componentes. En particular, los resultados en las bandas β y γ requieren una interpretación especialmente cautelosa cuando se observan cerca de los músculos del cuero cabelludo [18]. Por eso, la correcta identificación y tratamiento de este tipo de artefactos es una necesidad metodológica y no solo un detalle técnico.

En la Figura 1.3 se ilustra la morfología de los artefactos musculares de masticación y su efecto sobre la amplitud de la señal EEG. Durante las ráfagas EMG del tramo de masticación, los canales frontales presentan actividad de alta frecuencia superpuesta, especialmente visible en AF7, F7 y F8. En cambio, las deflexiones lentas de gran amplitud se distribuyen como el parpadeo del reposo, con mayor amplitud en Fp1, Fpz y Fp2, y coinciden con deflexiones del EOG (EXG1, no incluido en la figura). Por ello, podrían tener un origen ocular o de movimiento, y no solo muscular.

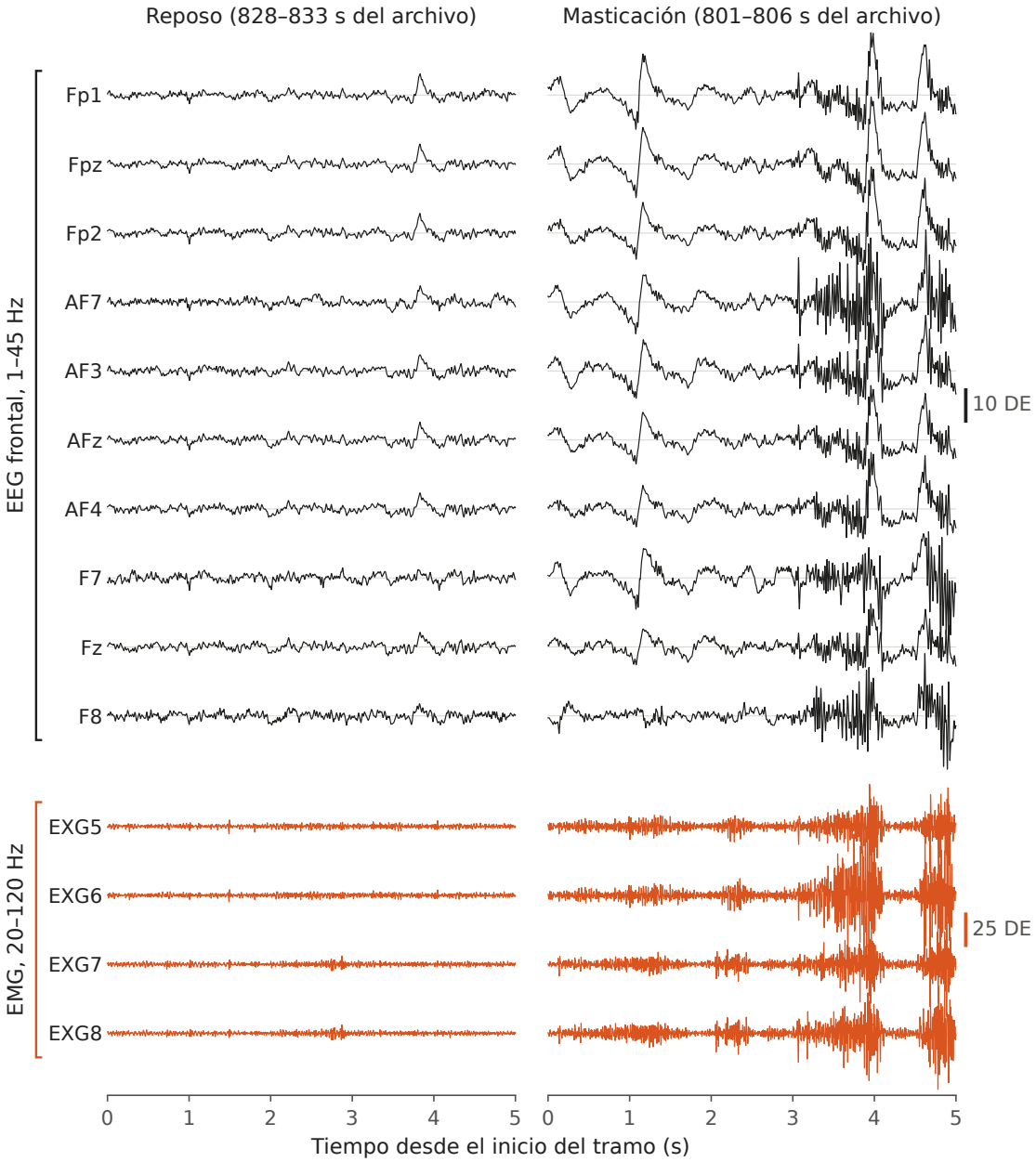


Figura 1.3: Diez canales EEG frontales y los cuatro canales EMG de los maseteros (EXG5 y EXG6, izquierdo; EXG7 y EXG8, derecho) de E1M1_R2, en 5 s de reposo (828–833 s, izquierda) y de masticación (801–806 s, derecha). Filtrado, referencia y normalización como en la Figura 1.1; misma escala vertical en ambas columnas, con barras de 10 DE (EEG) y 25 DE (EMG). En reposo, la deflexión común cercana a 3,8 s es compatible con un parpadeo. Elaboración propia a partir de registros de Sebastián Espinoza, laboratorio de neurociencia del AC3E (Sección 4.1).

CAPÍTULO

2

Estado del arte

2.1. Panorama de los métodos de limpieza

La literatura reporta múltiples estrategias para el tratamiento de artefactos en EEG [21, 22]. En términos generales, estas estrategias buscan atenuar los artefactos sin degradar la dinámica neuronal de interés. Sin embargo, ningún método es completo ni universal [22]. El más adecuado depende de la señal, del artefacto y de la proporción entre ambos [21], y para el artefacto muscular una revisión sugiere elegir la técnica según se disponga de uno, pocos o muchos canales [23].

Algunos enfoques detectan los datos contaminados y los retiran o reconstruyen. FASTER detecta mediante umbrales estadísticos los canales, épocas y componentes independientes atípicos y los elimina [24], y PREP estandariza las primeras etapas del procesamiento con estadísticos robustos e interpolación de los canales defectuosos, sin tratar la actividad muscular [25]. Los métodos de aprendizaje profundo, por su parte, informan resultados competitivos [26], pero requieren mucho EEG con una referencia limpia, que en EEGdenoiseNet y en IC-U-Net se obtuvo procesando EEG real con otros métodos [27–29]. Como esta tesis no dispone de tales datos y se centra en métodos que atenúan el artefacto sin descartar segmentos ni canales, estos enfoques no forman parte de la comparación. La evaluación de esas redes con mezclas de relación señal a ruido conocida es, sin embargo, un antecedente de la prueba semisintética (Sección 4.7).

Para elegir y evaluar un método conviene considerar también algunas observaciones sobre los músculos de la cara y la mandíbula. Entre cinco músculos faciales contraídos voluntariamente, el masetero produjo la mayor contaminación, con una distribución sobre el cuero cabelludo que dependió de cada sujeto [30]. Al apretar los dientes o morder, el mayor deterioro de la relación señal a ruido se observó en la banda gamma, allí de 32 a 100 Hz [8], y las unidades motoras de los músculos frontal y temporal interfieren en el EEG aun en reposo [31]. Como la contaminación cambia

de una persona a otra, en esta tesis cada configuración se evalúa en participantes que no intervinieron en su elección. Por otra parte, la regresión es problemática sin un buen canal de referencia para cada fuente de artefacto [32]; cuando el EMG se registra, puede usarse como referencia, principio del filtrado adaptativo y de iCanClean.

En esta tesis se evalúan cuatro métodos para atenuar los artefactos asociados a la masticación, que corresponden a enfoques distintos: la reconstrucción subespacial (ASR), el filtrado adaptativo (AF por bloques) y los métodos multivariados basados en la CCA con referencias (CCA con rezagos e iCanClean). iCanClean es uno de los métodos más recientes considerados: su formulación se describió en 2022 y se evaluó con un fantoma en 2023 [11, 12].

2.2. Separación ciega de fuentes

ICA modela el EEG como una mezcla lineal e instantánea de fuentes independientes [33]; las componentes identificadas como artefactos se eliminan y la señal se reconstruye con las restantes. ICA obtuvo resultados favorables frente a la regresión y al análisis de componentes principales (PCA) [32], sus variantes son las más usadas en estudios con movimiento [34], y el clasificador ICLabel automatiza la identificación de las componentes, entre ellas las musculares, con potencia de banda ancha elevada sobre 20–30 Hz [35]. Frente al EMG, ICA tiene limitaciones: supone fuentes espacialmente estacionarias y necesita un número de muestras de varias veces el cuadrado del número de canales [33, 36]. Los artefactos sin topografía estereotipada no se concentran en pocas componentes [37], y con contracciones del masetero la actividad miogénica no se eliminó por completo [2]. El ICA mejorado con wavelets (wICA) aplica un umbral wavelet a las componentes en lugar de rechazarlas, para no perder la actividad cerebral que contienen [38].

CCA busca pares de combinaciones lineales de dos conjuntos de variables cuya correlación sea máxima [39]. Como separación ciega de fuentes (BSS-CCA), estima las fuentes a partir del propio EEG, sin referencia externa, y elimina la actividad muscular, más parecida a un ruido, que queda en las componentes de menor autocorrelación [3, 6, 40]. En EEG ictal, aumentó del 62 % al 81 % la sensibilidad para localizar el inicio de las crisis [6], aunque sus componentes pueden mezclar varias fuentes [40]. Para pocos canales, se ha combinado con descomposiciones de la señal (EEMD-CCA, MEMD-CCA y SSA-CCA) [41–43], aunque EEMD-CCA solo se validó con artefactos de movimiento [41]; en registros clínicos largos, se ha combinado con wavelets y bosques aleatorios [44]. Con EEG simulado y actividad miogénica real, ningún enfoque dominó: el desempeño de ICA, CCA, la descomposición empírica en modos y las wavelets dependió del nivel de contaminación [45]. Este uso ciego difiere de la CCA con referencias de iCanClean y de la CCA con rezagos (Sección 3.4).

2.3. Reconstrucción de subespacios de artefactos (ASR)

Dentro de los métodos automáticos, ASR detecta y reconstruye los segmentos con un comportamiento anómalo respecto de una calibración basal [46], y su uso es cada vez más frecuente

en el EEG de vigilia [47]. Para ello, aplica PCA en ventanas deslizantes y elimina las componentes cuya amplitud supera un umbral; la señal se reconstruye a partir de las restantes. El umbral de cada componente se obtiene de la distribución de su raíz cuadrática media (RMS) en un segmento de calibración con poco artefacto y crece con el parámetro de tolerancia k (ecuación (3.12)) [46].

Su rendimiento depende de la selección de parámetros y de la calidad del segmento de calibración, lo que puede modificar de manera importante el nivel de limpieza [48, 49]. Sus autores sitúan k típicamente entre 5 y 7 [46]. En registros de conducción simulada, en cambio, el valor óptimo se ubicó entre 10 y 100 [48] y, en una evaluación posterior, entre 20 y 30, con más componentes oculares y musculares eliminadas que cerebrales [50]. Según un comentario editorial, esos valores elevados se explican en parte porque el ajuste gaussiano truncado que usa ASR subestima la desviación estándar [51]. Para la calibración se recomienda un minuto de reposo con pocos artefactos, de la misma sesión de registro [49].

2.4. Filtrado adaptativo con referencias

El filtrado adaptativo se basa en el principio de cancelación adaptativa de ruido [52]. Una entrada de referencia, que contiene ruido correlacionado de una forma desconocida con el de la entrada primaria, se filtra de manera adaptativa y se resta de esta; si la señal no está correlacionada con el ruido, minimizar la potencia de la salida equivale a estimar la señal por mínimos cuadrados. Los coeficientes pueden actualizarse con el algoritmo de mínimos cuadrados medios (LMS). El método cancela la interferencia con poca distorsión aunque las frecuencias se superpongan, siempre que exista una referencia adecuada; si esta contiene parte de la señal de interés, esa señal se cancela parcialmente [52]. Las referencias suelen provenir de sensores adicionales [5], aunque también pueden derivarse del propio EEG [12, 30].

En EEG, el filtrado adaptativo se ha aplicado a artefactos oculares con referencias de electrooculograma y mínimos cuadrados recursivos (RLS) [53], que con factor de olvido y longitud de filtro iguales a uno equivalen a la regresión en el dominio del tiempo [54], y como LMS multicanal con referencias de electrooculograma y de EMG [55]. Con contracciones faciales y referencias obtenidas de un único canal de EEG, el algoritmo adaptativo más adecuado cambió según la referencia representara la componente EEG o la EMG [30]. En una línea cercana, el EEG occipital se ha limpiado por regresión en ventanas de 1 s sobre canales auxiliares que incluyen el EMG mandibular [56]. El AF por bloques evaluado aquí usa como referencias el EMG de los maseteros y, en algunas configuraciones, su envolvente (Sección 3.3).

2.5. iCanClean

iCanClean usa CCA para identificar y eliminar los subespacios del registro contaminado más correlacionados con subespacios de las referencias de ruido: selecciona los pares canónicos cuya correlación al cuadrado alcanza un umbral definido por el usuario, estima su proyección sobre los

canales por mínimos cuadrados y la resta [12]. Sus componentes de ruido pueden formarse con las variables canónicas de los datos, con las de las referencias o con ambas [11]. Las referencias pueden ser electrodos de ruido de doble capa [12, 57], registros de EMG o seudorreferencias obtenidas filtrando el propio EEG [12]. Si comparten el electrodo de referencia del EEG, contienen también actividad cerebral, que la CCA puede eliminar por ser común a ambos conjuntos; por eso, los autores recomiendan rerreferenciarlas temporalmente a su propio promedio [12]. En esta tesis, las referencias principales son derivaciones bipolares de cada masetero, que no comparten esa referencia, y los canales monopolares referidos al promedio del EEG, que sí la comparten, se examinan como análisis de sensibilidad (Sección 4.2).

En un fantoma con fuentes cerebrales conocidas, iCanClean obtuvo el puntaje más alto frente a ASR, a una CCA automática y a un filtrado adaptativo en todas las condiciones con artefactos, en algunas por un margen estrecho. Con las fuentes de músculo facial, extraídas mediante ICA del EEG de un sujeto que masticaba, tragaba y bebía, ambos métodos con referencia usaron seudorreferencias, e iCanClean y ASR obtuvieron el mejor desempeño. Estas comparaciones provienen del grupo que desarrolló el método, que representó cada algoritmo con la configuración que maximizó una métrica calculable solo con fuentes cerebrales conocidas [12].

No existe una ventana ni un umbral únicos para iCanClean. Con 120 electrodos de ruido de doble capa durante la marcha, la configuración que más aumentó el número de componentes independientes de buena calidad fue una ventana de 4 s y un umbral de 0,65; el umbral óptimo disminuyó al reducir los canales de ruido, y los autores presentaron esos valores solo como un punto de partida y advirtieron que los electrodos de doble capa no capturan el artefacto muscular [58]. Otros trabajos usaron 2 s y 0,85 [59], valores que coinciden con los predeterminados de la implementación oficial [60], o 2 s y 0,95 con EMG de cuello para el artefacto muscular [61].

2.6. EEG durante la masticación

Para evitar los artefactos, varios estudios registraron el EEG solo antes y después de masticar, o detuvieron el registro durante la masticación [62]. Uno de ellos dejó al menos 60 s entre el fin de la masticación y el EEG analizado [9], y otro comparó potenciales evocados antes y después de ella [63]. Con EEG registrado durante la masticación, en cambio, se ha usado ICA para extraer la actividad alfa [64] y CCA con modelos de mezcla gaussiana [20]. Un trabajo aún no revisado por pares comparó durante la masticación iCanClean con seudorreferencia, ASR y un filtro de rechazo de banda, y obtuvo la mayor supresión al combinar iCanClean y ASR [7]. Estos trabajos difieren en el tipo de referencia, en los parámetros y en las métricas, lo que dificulta compararlos. En este contexto, una revisión reciente recomienda adquisiciones continuas de alta densidad, con al menos 64 canales, junto con EMG mandibular o sensores inerciales y métodos de sustracción de artefactos [62].

Enfoque	Referencia	Calibración o entrenamiento	Ventaja	Limitación
Detección y rechazo (FASTER, PREP) [24, 25]	No	Umbrales o estadísticos robustos del propio registro	Detecta de forma automática canales y segmentos contaminados	Retira o interpola datos; PREP no trata el EMG
ICA y clasificación de componentes [2, 32, 35–37]	No	Descomposición del propio registro; clasificador preentrenado	Separa el EMG espacialmente estereotipado	Separación incompleta del EMG; requiere muchos datos
BSS-CCA y variantes [3, 6, 40–43]	No	Descomposición de cada época o registro	El EMG queda en las componentes menos autocorrelacionadas	Las componentes pueden mezclar varias fuentes
ASR [46, 48–50]	No	Reposo con poco artefacto (se recomienda 1 min)	Automático y apto para tiempo real	Depende de la tolerancia k y de la calibración
Filtrado adaptativo [5, 12, 52]	Sí: sensores o seudorreferencias	Coeficientes que se adaptan durante el registro	Poca distorsión aunque las frecuencias se superpongan	Puede cancelar parte de la señal presente en la referencia
iCanClean [11, 12, 58]	Sí: electrodos de ruido, EMG o seudorreferencias	CCA en cada ventana; ventana y umbral fijados por el usuario	Elimina los subespacios del EEG más correlacionados con las referencias	Puede eliminar actividad cerebral común; ventana y umbral sin valor único
Aprendizaje profundo [27–29]	No, en las obras revisadas	Entrenamiento supervisado con referencia limpia	Una vez entrenado, no requiere canales de referencia	Requiere muchos datos con referencia limpia

Tabla 2.1: Comparación de los enfoques revisados para tratar artefactos en EEG. «Referencia» indica si el método usa señales relacionadas con el artefacto.

2.7. Síntesis y enfoque de esta tesis

La Tabla 2.1 resume estos enfoques: los métodos sin referencia dependen de supuestos sobre el EEG o el artefacto, y los demás, de que las referencias capten el artefacto y no la actividad cerebral.

En este contexto, esta tesis compara los cuatro métodos en registros reales de EEG con referencias EMG de los maseteros, en reposo y durante la masticación. El AF por bloques y la CCA con rezagos no reproducen un método publicado: son implementaciones propias basadas en los principios del filtrado adaptativo y de la CCA con referencias. Las aplicaciones revisadas de iCanClean usaron electrodos de doble capa, EMG de cuello o seudorreferencias [7, 12, 58, 61]; aquí se aplica con referencias de músculos masticatorios y en las mismas condiciones que los demás algoritmos. Esta comparación bajo condiciones comunes constituye una de las principales contribuciones de esta tesis. La otra es la variante extendida propia (Sección 3.6).

La evaluación parte de un preprocesamiento común y combina la inspección visual de las señales y de sus espectros con métricas de similitud, de cambio de amplitud y de potencia por banda (Sección 4.5); el efecto de los parámetros se examina con una grilla de configuraciones.

CAPÍTULO

3

Marco teórico

Este capítulo define el problema de limpieza de artefactos musculares en el EEG y fija la notación de la tesis. El objetivo es establecer una base común que permita comparar métodos con supuestos diferentes. La comparación se apoya en una misma señal observada, un conjunto explícito de canales de referencia y una separación clara entre calibración, estimación del artefacto y obtención de la señal corregida.

3.1. Notación, preprocesamiento común y principios de comparación

3.1.1. Convención y modelo de observación

Se adopta la siguiente convención. La señal EEG observada se representa por

$$\boldsymbol{X} \in \mathbb{R}^{T \times C}, \tag{3.1}$$

donde T es el número de muestras y C el número de canales EEG. Cada fila corresponde a una muestra y cada columna a un canal. El índice de muestra se denota por $n = 1, \ldots, T$; la fila n se escribe $\boldsymbol{x}[n] \in \mathbb{R}^{1 \times C}$ y el valor del canal c en esa muestra, $x_c[n]$. Cuando un registro contiene discontinuidades, cada parte comprendida entre ellas se denomina pieza continua y se procesa por separado (Sección 4.3). La observación se modela como

$$\boldsymbol{X} = \boldsymbol{S} + \boldsymbol{A} + \boldsymbol{\varepsilon}, \tag{3.2}$$

donde $\boldsymbol{S}$ es la actividad neuronal de interés, $\boldsymbol{A}$ es la contribución del artefacto y $\boldsymbol{\varepsilon}$ agrupa el ruido residual y los errores de medición.

Junto con el EEG se registran canales EMG, con los que se forman las referencias del artefacto. Filtradas en la misma banda que el EEG, se reúnen en la matriz

$$\boldsymbol{R} \in \mathbb{R}^{T \times Q_{\mathrm{EMG}}}, \tag{3.3}$$

donde Q_{EMG} es el número de referencias: $Q_{\mathrm{EMG}} = 2$ con una derivación bipolar por masetero, en la configuración principal, y $Q_{\mathrm{EMG}} = 4$ con los canales monopolares de los análisis de sensibilidad (Sección 4.2). Su fila n se denota $\boldsymbol{r}[n]$ y el valor de la referencia q, $r_q[n]$. La salida de cualquier método se denomina $\boldsymbol{X}_{\mathrm{clean}}$ y conserva la dimensión de $\boldsymbol{X}$; cuando un método estima explícitamente la contribución del artefacto, esa estimación se denota $\widehat{\boldsymbol{A}} \in \mathbb{R}^{T \times C}$, con valores $\widehat{a}_c[n]$.

La ecuación (3.2) es un modelo conceptual, no una identidad observable. En la práctica, $\boldsymbol{S}$ y $\boldsymbol{A}$ no se conocen por separado. Por esa razón, la evaluación combina métricas cuantitativas, inspección de segmentos de reposo y masticación, y una descripción explícita de los supuestos de cada algoritmo. La Figura 3.1 resume este problema común.

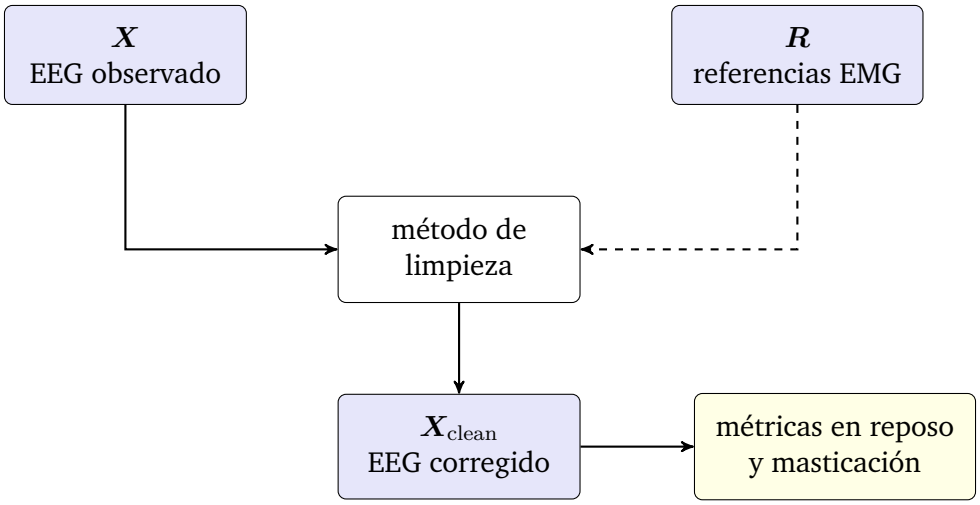


Figura 3.1: Problema común de la tesis. La señal EEG observada y, en los métodos que las utilizan (línea discontinua), las referencias EMG entran al método de limpieza. La señal corregida se evalúa con el mismo procedimiento en reposo y en masticación para todos los métodos.

Varios métodos estiman parámetros con un segmento de calibración, que se identifica con el subíndice cal. Sean $\bar{\boldsymbol{x}}_{\mathrm{cal}} \in \mathbb{R}^{1 \times C}$ las medias por canal de ese segmento y $\boldsymbol{\sigma}_{\mathrm{cal}} \in \mathbb{R}^{1 \times C}$ sus desviaciones estándar. El subíndice z indica una matriz estandarizada por columna con esos estadísticos:

$$\boldsymbol{X}_z = (\boldsymbol{X} - \mathbf{1}\bar{\boldsymbol{x}}_{\mathrm{cal}})\,\mathrm{diag}(\boldsymbol{\sigma}_{\mathrm{cal}})^{-1}, \tag{3.4}$$

donde $\mathbf{1}$ es un vector columna de unos. La misma operación se aplica a las referencias con sus propios estadísticos de calibración. El subíndice MC indica, en cambio, una matriz centrada por columna con las medias del mismo segmento: $\boldsymbol{X}_{\mathrm{MC}} = \boldsymbol{X} - \mathbf{1}\bar{\boldsymbol{x}}$ y $\boldsymbol{R}_{\mathrm{MC}} = \boldsymbol{R} - \mathbf{1}\bar{\boldsymbol{r}}$, donde $\bar{\boldsymbol{x}}$ y $\bar{\boldsymbol{r}}$ son vectores fila de medias.

3.1.2. Preprocesamiento común

Como la preparación de los datos influye en el desempeño de los métodos [25], el preprocesamiento común reduce las diferencias atribuibles a esa preparación y reserva la comparación para el algoritmo de limpieza. El tipado asigna a cada canal su tipo (EEG, EOG o EMG); la rerreferencia lo expresa respecto de una referencia, como el promedio de los canales EEG [65]; el filtrado conserva la banda de 1–45 Hz, y el montaje asigna a cada electrodo su posición. Al EEG no se le aplica un filtro de rechazo de banda (*notch*), porque la red eléctrica queda fuera de esa banda. El mismo segmento preprocesado $\boldsymbol{X}$ se entrega a todos los métodos, y la evaluación se resume en un conjunto fijo de canales (Secciones 4.2 y 4.5).

3.1.3. Representaciones de las referencias

Para un entero j, la copia desplazada $\boldsymbol{R}^{(j)} \in \mathbb{R}^{T\times Q_{\mathrm{EMG}}}$ tiene por fila

$$\boldsymbol{r}^{(j)}[n] = \begin{cases} \boldsymbol{r}[n-j], & 1 \le n-j \le T, \\ \boldsymbol{0}, & \text{en otro caso.} \end{cases} \tag{3.5}$$

Con $j > 0$, la fila n contiene un valor anterior de la referencia (rezago); con $j < 0$, uno siguiente (adelanto). La matriz de referencias con rezagos simétricos es

$$\boldsymbol{R}_{\mathrm{lag}} = \left[\boldsymbol{R}^{(-L)}, \ldots, \boldsymbol{R}^{(0)}, \ldots, \boldsymbol{R}^{(L)}\right] \in \mathbb{R}^{T\times(2L+1)Q_{\mathrm{EMG}}}, \tag{3.6}$$

donde L es el rezago máximo, en muestras. Con $L = 0$, $\boldsymbol{R}_{\mathrm{lag}} = \boldsymbol{R}$. Los rezagos representan desfases entre las referencias y su manifestación en el EEG. Los desplazamientos se calculan dentro de cada pieza continua, y las posiciones que quedarían fuera de ella se completan con ceros.

La envolvente, por su parte, resume la evolución lenta de la amplitud de cada referencia. Se calcula a partir del EMG filtrado en 20–120 Hz con rechazo de banda en 50 y 100 Hz, una señal distinta de $\boldsymbol{R}$ que se denota $\breve{r}_q[n]$:

$$r_q^{\mathrm{env}}[n] = \left(h_{\mathrm{EEG}} * \left[\mathcal{H}_{\mathrm{LP}}\left\{|\breve{r}_q|\right\}\right]_{\mathrm{MC}}\right)[n], \qquad q = 1, \ldots, Q_{\mathrm{EMG}}, \tag{3.7}$$

donde el valor absoluto se suaviza con un filtro pasa-bajos $\mathcal{H}_{\mathrm{LP}}\{\cdot\}$, de frecuencia de corte f_{env}, aplicado hacia adelante y hacia atrás para no introducir desfase; $[\cdot]_{\mathrm{MC}}$ resta la media en la pieza, y h_{EEG} es el filtro de 1–45 Hz del EEG (Sección 4.2).

La envolvente sigue así la intensidad de la actividad muscular, que puede variar con los ciclos de masticación, y se incluye porque parte del artefacto en las bandas bajas del EEG podría seguir esa intensidad y no la forma de onda del EMG. Sus columnas forman la matriz $\boldsymbol{R}_{\mathrm{env}} \in \mathbb{R}^{T\times Q_{\mathrm{EMG}}}$ y, cuando un método la utiliza, la referencia ampliada es

$$\tilde{\boldsymbol{R}} = [\boldsymbol{R}, \boldsymbol{R}_{\mathrm{env}}] \in \mathbb{R}^{T\times\tilde{Q}}, \qquad \tilde{Q} = 2Q_{\mathrm{EMG}}; \tag{3.8}$$

si no la utiliza, $\tilde{\boldsymbol{R}} = \boldsymbol{R}$ y $\tilde{Q} = Q_{\mathrm{EMG}}$. Los desplazamientos de las ecuaciones (3.5) y (3.6) se aplican de la misma manera a $\tilde{\boldsymbol{R}}$, con lo que se obtienen $\tilde{\boldsymbol{R}}^{(j)}$ y la matriz de rezagos $\tilde{\boldsymbol{R}}_{\mathrm{lag}} \in \mathbb{R}^{T\times(2L+1)\tilde{Q}}$.

Como los adelantos, la envolvente, el filtrado de fase cero y la anticipación de ASR usan muestras siguientes a la corregida, el procesamiento es no causal, lo que es compatible con el análisis fuera de línea.

3.1.4. Principios de comparación

La adaptación no tiene el mismo significado en todos los métodos. ASR estima una referencia estadística a partir de un segmento con poco artefacto [46], aquí de reposo. El AF, la CCA con rezagos y la variante extendida estiman sus estadísticos de normalización y sus parámetros con un segmento de calibración de masticación; después, el AF sigue actualizando sus pesos, y la CCA con rezagos reestima la selección y la proyección en el segmento que limpia. iCanClean, en cambio, se estima en cada ventana que se limpia [11]. Por lo tanto, no existe una única etapa de entrenamiento común para todos los algoritmos (Sección 4.3).

La salida se compara mediante medidas de similitud con el EEG observado y de cambio de amplitud, calculadas por separado en reposo y en masticación, que se definen en la Sección 4.5.

3.2. Reconstrucción de subespacios de artefactos (ASR)

3.2.1. Supuesto y objetivo

ASR modela la dinámica multicanal del EEG en una condición de referencia relativamente limpia [46, 66]. Su supuesto principal es que la actividad fisiológica esperada ocupa una estructura estadística que puede estimarse mediante la covarianza de un segmento de calibración. Una contaminación transitoria de gran amplitud modifica la energía de ciertas direcciones del espacio de canales. Por lo tanto, ASR detecta esas desviaciones y reconstruye las componentes afectadas a partir de las direcciones restantes, sin descartar la ventana temporal completa [46].

El método se ajusta a registros continuos porque decide la corrección en ventanas cortas, típicamente de 0,5 s [46, 49].

3.2.2. Representación matricial

Sea $\boldsymbol{X}_{\mathrm{cal}} \in \mathbb{R}^{T_{\mathrm{cal}}\times C}$ el segmento de calibración, tomado del reposo (Sección 4.3). Su covarianza muestral es

$$\boldsymbol{\Sigma}_{\mathrm{cal}} = \frac{1}{T_{\mathrm{cal}} - 1} \left(\boldsymbol{X}_{\mathrm{cal}} - \mathbf{1}\bar{\boldsymbol{x}}_{\mathrm{cal}}\right)^{\mathsf{T}} \left(\boldsymbol{X}_{\mathrm{cal}} - \mathbf{1}\bar{\boldsymbol{x}}_{\mathrm{cal}}\right), \tag{3.9}$$

donde $T_{\mathrm{cal}} > 1$ es el número de muestras de calibración. La matriz $\boldsymbol{\Sigma}_{\mathrm{cal}}$, de dimensión $C \times C$, contiene la varianza de cada canal en su diagonal y la covarianza entre pares de canales fuera de ella.

Como la covarianza muestral es sensible a los tramos de gran amplitud, la implementación descarta primero las ventanas de calibración con amplitudes atípicas (Sección 4.4), y ASR usa una versión robusta de esta referencia [46, 49]. Antes de estimarla, la señal se pondera espectralmente con un filtro que realza las frecuencias típicas de los artefactos y atenúa ritmos cerebrales de gran amplitud [46]. La señal ponderada, $\check{\boldsymbol{X}}_{\mathrm{cal}}$, se divide en bloques consecutivos y, en cada bloque β, se calcula la matriz de segundos momentos $\boldsymbol{\Sigma}_\beta = \check{\boldsymbol{X}}_\beta^{\mathsf{T}} \check{\boldsymbol{X}}_\beta / T_\beta$, donde $\check{\boldsymbol{X}}_\beta$ contiene las T_β muestras del bloque [49, 67]. La referencia robusta es la mediana geométrica de esas matrices [46, 49]:

$$\boldsymbol{\Sigma}_* = \arg\min_{\boldsymbol{\Sigma}} \sum_\beta \|\boldsymbol{\Sigma}_\beta - \boldsymbol{\Sigma}\|_F. \tag{3.10}$$

Si la señal no se ponderara, estuviera centrada y los bloques tuvieran igual longitud, el promedio de las matrices $\boldsymbol{\Sigma}_\beta$ coincidiría con $\boldsymbol{\Sigma}_{\mathrm{cal}}$ salvo el factor $(T_{\mathrm{cal}} - 1)/T_{\mathrm{cal}}$; la mediana geométrica reduce, en cambio, la influencia de los bloques con artefactos. La raíz cuadrada simétrica de la referencia robusta,

$$\boldsymbol{M} = \boldsymbol{\Sigma}_*^{1/2}, \qquad \boldsymbol{M}\boldsymbol{M}^{\mathsf{T}} = \boldsymbol{\Sigma}_*, \tag{3.11}$$

interviene directamente en la reconstrucción de la ecuación (3.15) [46]. Su inversa, $\boldsymbol{M}^{-1}$, es la transformación de blanqueamiento asociada a la calibración, ya que $\boldsymbol{M}^{-1}\boldsymbol{\Sigma}_*\boldsymbol{M}^{-1} = \boldsymbol{I}$.

Los umbrales se derivan de la misma calibración. Sea $\boldsymbol{\Phi}_{\mathrm{cal}} \in \mathbb{R}^{C\times C}$ la matriz cuyas columnas son los vectores propios de $\boldsymbol{M}$, que coinciden con los de $\boldsymbol{\Sigma}_*$. La proyección de la señal ponderada sobre cada vector propio es una componente cuya amplitud RMS se calcula en ventanas cortas y solapadas. A la distribución de esas amplitudes se ajustan parámetros robustos de posición y escala, mediante una distribución truncada que tolera una fracción de ventanas con artefactos; esos parámetros se reúnen en los vectores $\boldsymbol{\mu}_{\mathrm{ASR}}$ y $\boldsymbol{\sigma}_{\mathrm{ASR}}$ [46]. Con ellos se construye la matriz de umbrales

$$\boldsymbol{\Theta} = \mathrm{diag}\left(\boldsymbol{\mu}_{\mathrm{ASR}} + k\,\boldsymbol{\sigma}_{\mathrm{ASR}}\right) \boldsymbol{\Phi}_{\mathrm{cal}}^{\mathsf{T}}, \tag{3.12}$$

donde k es la tolerancia frente a desviaciones de la referencia [46, 49]. Para una dirección unitaria $\boldsymbol{\phi}$ del espacio de canales, el umbral es $\|\boldsymbol{\Theta}\boldsymbol{\phi}\|_2^2$ [46]. Esta cantidad está en unidades de varianza y, por lo tanto, puede compararse con los valores propios de cada ventana.

En cada ventana w se calculan la matriz local de segundos momentos de la señal ponderada, $\boldsymbol{\Sigma}_w$, y su descomposición en valores y vectores propios:

$$\boldsymbol{\Sigma}_w = \boldsymbol{\Phi}_w\, \mathrm{diag}(\boldsymbol{d}_w) \boldsymbol{\Phi}_w^{\mathsf{T}}. \tag{3.13}$$

Las columnas $\boldsymbol{\phi}_{w,j}$ de $\boldsymbol{\Phi}_w$ son las direcciones principales de la ventana, y los valores propios $d_{w,j}$, ordenados de menor a mayor, son las varianzas de la señal en esas direcciones. Una dirección se

conserva si su varianza es menor que el umbral de la calibración o si su índice es menor que $C - m$, es decir, si está entre las de menor varianza:

$$\delta_{w,j} = \left[d_{w,j} < \|\boldsymbol{\Theta}\boldsymbol{\phi}_{w,j}\|_2^2 \right] \ \vee \ [j < C - m], \qquad j = 1, \ldots, C, \tag{3.14}$$

donde $[\cdot]$ vale 1 si la condición se cumple y 0 en otro caso. El parámetro m se obtiene de una fracción f_{ASR} del número de canales [67, 68]. Como j recorre los valores $1, \ldots, C$, las $C - m - 1$ direcciones de menor varianza se conservan siempre, de modo que en cada ventana pueden reconstruirse hasta $m + 1$ direcciones (Apéndice A.2).

Con $\boldsymbol{D}_w = \mathrm{diag}(\boldsymbol{\delta}_w)$, la matriz de reconstrucción de la ventana y la ventana reconstruida son

$$\boldsymbol{G}_w = \boldsymbol{M} \left(\boldsymbol{D}_w \boldsymbol{\Phi}_w^{\mathsf{T}} \boldsymbol{M} \right)^{+} \boldsymbol{\Phi}_w^{\mathsf{T}}, \qquad \widehat{\boldsymbol{X}}_w = \boldsymbol{X}_w \boldsymbol{G}_w^{\mathsf{T}}, \tag{3.15}$$

donde $(\cdot)^+$ denota la pseudoinversa de Moore–Penrose y $\boldsymbol{X}_w$ reúne las muestras de la ventana sin ponderar, con la convención de muestras por filas [46]. La matriz $\boldsymbol{D}_w \boldsymbol{\Phi}_w^{\mathsf{T}} \boldsymbol{M}$ conserva las filas de las direcciones aceptadas y anula las demás. Su pseudoinversa reconstruye la actividad de las direcciones rechazadas como una combinación lineal de las conservadas, de acuerdo con la estructura de covarianza de la calibración contenida en $\boldsymbol{M}$ [46, 66]. Si todas las direcciones se conservan, $\boldsymbol{G}_w = \boldsymbol{I}$ y la ventana no se modifica. La ponderación espectral interviene, a través de $\boldsymbol{M}$, $\boldsymbol{\Theta}$ y $\boldsymbol{\Sigma}_w$, en qué direcciones se rechazan y en cómo se reconstruyen, pero no se aplica a la señal reconstruida, de modo que no filtra directamente la salida [46].

La matriz $\boldsymbol{G}_w$ se recalcula en puntos de actualización n_w separados por un número fijo de muestras. Entre dos actualizaciones consecutivas, cada muestra se corrige con una mezcla de las dos matrices vecinas [46]:

$$\boldsymbol{x}_{\mathrm{clean}}[n] = \xi_n \, \boldsymbol{x}[n] \boldsymbol{G}_w^{\mathsf{T}} + (1 - \xi_n) \, \boldsymbol{x}[n] \boldsymbol{G}_{w-1}^{\mathsf{T}}, \qquad \xi_n = \frac{1}{2} \left(1 - \cos \frac{\pi \, (n - n_{w-1})}{n_w - n_{w-1}} \right), \tag{3.16}$$

para $n_{w-1} < n \leq n_w$. El peso ξ_n crece de forma gradual desde casi 0 hasta 1, de modo que la corrección pasa de la matriz anterior a la nueva sin saltos. Las filas $\boldsymbol{x}_{\mathrm{clean}}[n]$ forman la señal corregida $\boldsymbol{X}_{\mathrm{clean}}$. La Figura 3.2 resume el procedimiento.

3.2.3. Parámetros e interpretación

El parámetro k de la ecuación (3.12) establece la tolerancia frente a desviaciones de la referencia. Un valor menor tiende a corregir más direcciones y aumenta el riesgo de sobrecorrección; un valor mayor conserva más variabilidad, pero puede dejar contaminación residual. La duración de la ventana determina el compromiso entre adaptación local y estabilidad de la estimación. La fracción f_{ASR} limita el número de direcciones que pueden reconstruirse en una ventana, y el intervalo de actualización fija cada cuántas muestras se recalcula $\boldsymbol{G}_w$. La anticipación permite que la ventana de estadísticas incluya muestras siguientes a la que se corrige; en una aplicación en línea, equivale a

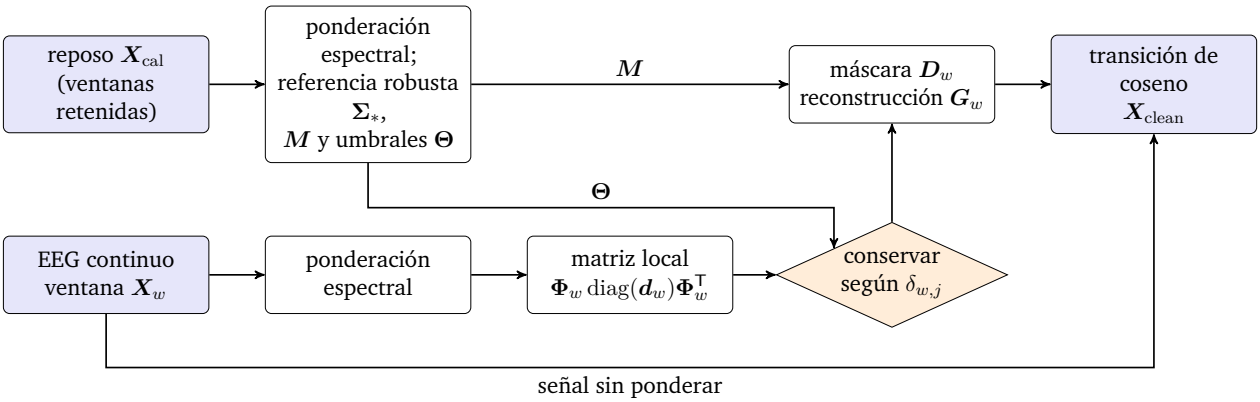


Figura 3.2: Flujo de ASR. La calibración en reposo, restringida a sus ventanas sin amplitudes atípicas, define la referencia robusta $\boldsymbol{\Sigma}_*$, su raíz $\boldsymbol{M}$ y la matriz de umbrales $\boldsymbol{\Theta}$. La ponderación espectral se aplica a la señal con que se estiman estas matrices y la matriz local de cada ventana; la reconstrucción y la transición de coseno se aplican a la señal sin ponderar. En cada ventana, la regla $\delta_{w,j}$ determina las direcciones conservadas, $\boldsymbol{G}_w$ reconstruye las rechazadas y la transición de coseno entre matrices sucesivas forma $\boldsymbol{X}_{\text{clean}}$.

retrasar la salida en ese mismo intervalo [67]. Así, una salida de ASR solo queda definida junto con estos parámetros (Sección 4.4).

3.2.4. Limitaciones

ASR depende de que el segmento de calibración represente de forma razonable la variabilidad fisiológica que se desea conservar [46, 49]. Una calibración demasiado corta o contaminada puede producir umbrales poco representativos. Asimismo, la reconstrucción subespacial puede atenuar actividad neuronal de gran amplitud si esta se parece estadísticamente a una desviación artefactual. Por ejemplo, con los parámetros recomendados para la vigilia, ASR eliminó ondas lentas del EEG de sueño [47]. Un comentario reciente advierte, además, que un filtro espacial como el de ASR descompone con dificultad las señales de alta frecuencia [51], una advertencia pertinente para el artefacto muscular.

3.3. Filtrado adaptativo (AF)

3.3.1. Modelo de señal

AF utiliza una referencia correlacionada con el artefacto y, en lo posible, poco correlacionada con la actividad neuronal [52]. Para un canal EEG, la observación se escribe como

$$x_c[n] = s_c[n] + a_c[n], \tag{3.17}$$

donde $x_c[n]$ es la observación, $s_c[n]$ la señal neuronal y $a_c[n]$ el artefacto, es decir, la columna c del modelo de la ecuación (3.2), sin el término de ruido. El filtro reúne valores actuales y pasados de las referencias en un vector fila $\boldsymbol{\psi}[n]$ y produce una estimación del artefacto

$$\widehat{a}_c[n] = \boldsymbol{\psi}[n]\, \boldsymbol{w}_c[n], \tag{3.18}$$

donde $\boldsymbol{w}_c[n]$ es el vector de pesos del canal c. La señal corregida es el residual

$$e_c[n] = x_c[n] - \widehat{a}_c[n]. \tag{3.19}$$

El residual es la parte de la observación que el filtro no predice con sus regresores.

3.3.2. Actualización normalizada por bloques

La regla clásica de mínimos cuadrados medios (LMS) actualiza los pesos en cada muestra, en la dirección que reduce el cuadrado del residual [52]:

$$\boldsymbol{w}_c[n+1] = \boldsymbol{w}_c[n] + \mu\, e_c[n]\, \boldsymbol{\psi}[n]^{\mathsf{T}}, \tag{3.20}$$

donde μ es el paso de adaptación; el factor 2 de la formulación original se incorpora en μ. El paso controla la estabilidad y la velocidad de convergencia: si es excesivamente pequeño, la convergencia es lenta, y si es demasiado grande, los pesos pueden oscilar o volverse inestables [52].

Se utiliza aquí una versión normalizada por bloques, el AF por bloques, con cuatro modificaciones respecto de la ecuación (3.20): el objetivo y las referencias se estandarizan con estadísticos de la calibración; el paso se divide por la energía de los regresores; los pesos se actualizan una vez por bloque de B muestras, con una fuga ν; y la predicción se centra en cada pieza antes de restarse. El objetivo normalizado del canal c es

$$\tilde{x}_c[n] = \frac{x_c[n] - \bar{x}_{\mathrm{cal},c}}{\sigma_{\mathrm{cal},c}}, \tag{3.21}$$

el elemento (n, c) de $\boldsymbol{X}_z$ en la ecuación (3.4). En el AF, la tilde sobre el objetivo, la predicción y el residual indica magnitudes normalizadas; sobre $\tilde{\boldsymbol{R}}$ y $\tilde{Q}$, la referencia ampliada. Las referencias $\tilde{\boldsymbol{R}}$, con o sin envolvente, se estandarizan del mismo modo y forman el vector de regresores

$$\boldsymbol{\psi}[n] = [\tilde{\boldsymbol{r}}_z[n],\, \tilde{\boldsymbol{r}}_z[n-1],\, \ldots,\, \tilde{\boldsymbol{r}}_z[n - L_{\mathrm{AF}} + 1],\, 1] \in \mathbb{R}^{1\times(\tilde{Q}L_{\mathrm{AF}}+1)}, \tag{3.22}$$

donde $\tilde{\boldsymbol{r}}_z[n]$ es la fila n de $\tilde{\boldsymbol{R}}_z$. El vector contiene los rezagos causales $0, \ldots, L_{\mathrm{AF}} - 1$ de cada referencia y un término de sesgo constante, que se omite si no se utiliza; L_{AF} es la memoria del filtro, en muestras, y los valores anteriores al inicio de cada pieza continua se toman como cero.

El bloque b reúne N_b muestras consecutivas, con $N_b = B$ salvo, posiblemente, en el último bloque de cada pieza. Sean $\boldsymbol{\Psi}_b \in \mathbb{R}^{N_b\times(\tilde{Q}L_{\mathrm{AF}}+1)}$ las filas $\boldsymbol{\psi}[n]$ del bloque, $\tilde{\boldsymbol{X}}_b \in \mathbb{R}^{N_b\times C}$ los objetivos

normalizados y $\boldsymbol{W}_b$ la matriz de pesos vigente al inicio del bloque, cuya columna c es $\boldsymbol{w}_c$. La predicción se calcula a priori, con los pesos anteriores a la actualización:

$$\tilde{\boldsymbol{A}}_b = \boldsymbol{\Psi}_b \boldsymbol{W}_b, \qquad \boldsymbol{E}_b = \tilde{\boldsymbol{X}}_b - \tilde{\boldsymbol{A}}_b. \tag{3.23}$$

Los elementos de $\tilde{\boldsymbol{A}}_b$ y $\boldsymbol{E}_b$ son la predicción normalizada $\tilde{a}_c[n]$ y el residual normalizado $\tilde{e}_c[n] = \tilde{x}_c[n] - \tilde{a}_c[n]$. La energía media de los regresores en el bloque y la actualización de los pesos son

$$\eta_b = \frac{1}{N_b} \|\boldsymbol{\Psi}_b\|_F^2 + 10^{-6}, \qquad \boldsymbol{W}_{b+1} = (1-\nu)\, \boldsymbol{W}_b + \frac{\mu}{\eta_b} \frac{1}{N_b} \boldsymbol{\Psi}_b^{\mathsf{T}} \boldsymbol{E}_b, \tag{3.24}$$

donde la constante 10^{-6} evita la división por cero. La división por η_b hace que el tamaño efectivo del paso no dependa de la escala de los regresores, y el factor $(1-\nu)$ reduce los pesos cuando el residual no los sostiene.

En la calibración, los pesos parten de cero y se actualizan en una pasada por el segmento. En cada pieza de evaluación, el filtro parte de una copia de los pesos calibrados y continúa adaptándolos para seguir los cambios de la relación entre referencias y EEG. Recorrida la pieza, la predicción se centra restando su media en ella, $\bar{\tilde{a}}_c$, y se devuelve a las unidades del EEG:

$$\widehat{a}_c[n] = \sigma_{\mathrm{cal},c}(\tilde{a}_c[n] - \bar{\tilde{a}}_c), \qquad x_{\mathrm{clean},c}[n] = x_c[n] - \widehat{a}_c[n]. \tag{3.25}$$

Con referencias nulas, el término de sesgo es el único regresor activo y sigue la media local de $\tilde{x}_c[n]$, de modo que su resta actúa como un filtro pasa-altos sobre el propio EEG. Por ello, su inclusión se decide con una regla aplicada a los datos de desarrollo (Sección 4.3 y Apéndice A.5).

Con estadísticas estables, la fuga equivale aproximadamente a una regularización ridge de intensidad $\nu\eta/\mu$, donde η, la energía media de los regresores, crece con la memoria y con la envolvente; así, el paso, la memoria y la envolvente modifican a la vez la velocidad de adaptación y el grado de regularización. La predicción se resta completa, y los valores de los parámetros se presentan en la Sección 4.4. La Figura 3.3 resume el procedimiento.

3.3.3. Limitaciones

AF requiere una referencia informativa [52]. Si la referencia no sigue el artefacto, la resta puede ser débil. Si, en cambio, la referencia contiene actividad neuronal, la cancelación puede retirar también parte de la señal de interés [52]. El filtro tampoco asegura por sí mismo que el residual sea fisiológicamente válido. La referencia, la memoria, el paso de adaptación y el criterio de evaluación se informan, en consecuencia, de manera conjunta.

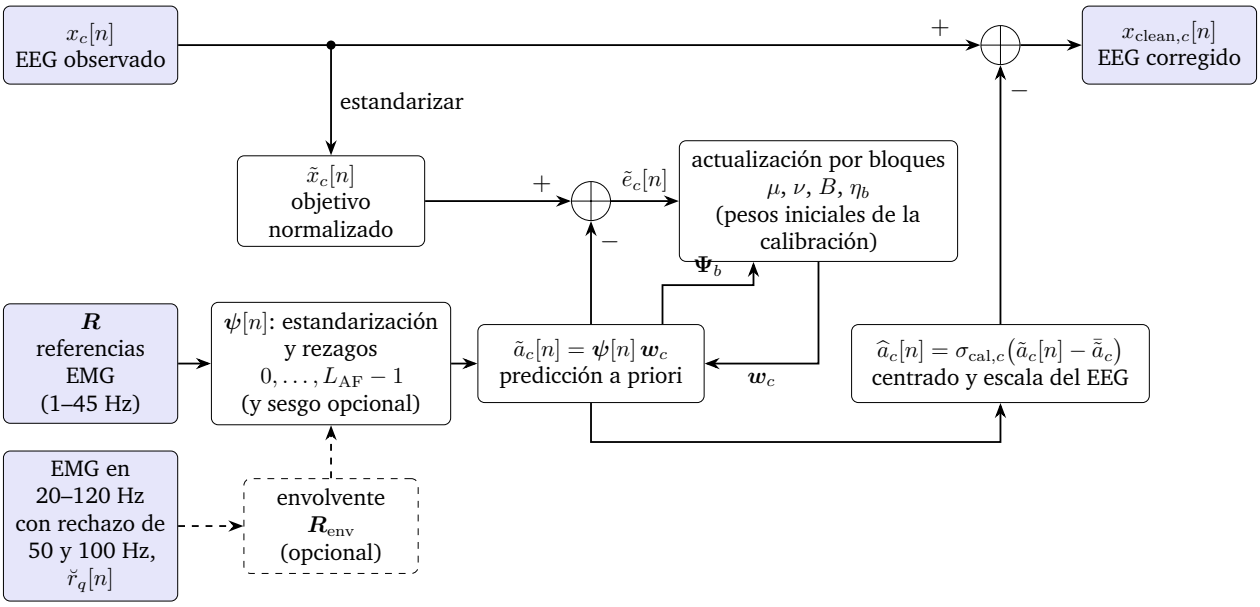


Figura 3.3: AF por bloques. El objetivo y las referencias se estandarizan con la calibración, con la que también se obtienen los pesos iniciales. La predicción a priori $\tilde{a}_c[n]$ se compara con $\tilde{x}_c[n]$; el residual normalizado y los regresores del bloque, $\boldsymbol{\Psi}_b$, actualizan los pesos cada B muestras, y la predicción centrada y reescalada, $\widehat{a}_c[n]$, se resta del EEG observado. La envolvente (línea discontinua) es opcional y se calcula a partir del EMG filtrado en 20–120 Hz con rechazo de banda en 50 y 100 Hz.

3.4. Análisis de correlación canónica (CCA)

CCA relaciona dos conjuntos de variables medidos simultáneamente; en esta aplicación, el EEG y las referencias. Busca pares de combinaciones lineales, una de cada conjunto, con la mayor correlación posible. Esas combinaciones se denominan variables canónicas y sus correlaciones, correlaciones canónicas [39]. A diferencia de la separación ciega de fuentes (Sección 2.2), aquí el segundo conjunto son las referencias EMG, como en iCanClean.

Sean $\boldsymbol{X}_{\text{MC}}$ y $\boldsymbol{R}_{\text{MC}}$ las matrices centradas. Sus matrices de covarianza son $\boldsymbol{\Sigma}_{XX} = \boldsymbol{X}_{\text{MC}}^{\mathsf{T}}\boldsymbol{X}_{\text{MC}}/(T-1)$, $\boldsymbol{\Sigma}_{RR} = \boldsymbol{R}_{\text{MC}}^{\mathsf{T}}\boldsymbol{R}_{\text{MC}}/(T-1)$ y $\boldsymbol{\Sigma}_{XR} = \boldsymbol{X}_{\text{MC}}^{\mathsf{T}}\boldsymbol{R}_{\text{MC}}/(T-1)$. Un par canónico está formado por las combinaciones $\boldsymbol{X}_{\text{MC}}\boldsymbol{\omega}_X$ y $\boldsymbol{R}_{\text{MC}}\boldsymbol{\omega}_R$. El primer par se obtiene resolviendo

$$\max_{\boldsymbol{\omega}_X, \boldsymbol{\omega}_R} \ \rho = \boldsymbol{\omega}_X^{\mathsf{T}}\boldsymbol{\Sigma}_{XR}\,\boldsymbol{\omega}_R \quad \text{sujeto a} \quad \boldsymbol{\omega}_X^{\mathsf{T}}\boldsymbol{\Sigma}_{XX}\,\boldsymbol{\omega}_X = 1, \quad \boldsymbol{\omega}_R^{\mathsf{T}}\boldsymbol{\Sigma}_{RR}\,\boldsymbol{\omega}_R = 1. \tag{3.26}$$

Las restricciones fijan en 1 la varianza de las dos combinaciones; de ahí que la función objetivo sea directamente la correlación entre ellas. Los pares siguientes cumplen la misma condición y no están correlacionados con los anteriores. Al reunir en columnas los pesos de todos los pares se obtienen las matrices de pesos canónicos $\boldsymbol{\Omega}_X$ y $\boldsymbol{\Omega}_R$, y las variables canónicas

$$\boldsymbol{U} = \boldsymbol{X}_{\text{MC}}\boldsymbol{\Omega}_X, \qquad \boldsymbol{V} = \boldsymbol{R}_{\text{MC}}\boldsymbol{\Omega}_R. \tag{3.27}$$

El par i está formado por las columnas $\boldsymbol{U}_{:,i}$ y $\boldsymbol{V}_{:,i}$, y su correlación ρ_i es la i-ésima correlación canónica; las correlaciones se numeran de mayor a menor.

3.4.1. Selección y sustracción

La CCA con rezagos aplica estas definiciones al EEG y a las referencias con rezagos $\boldsymbol{R}_{\text{lag}}$ de la ecuación (3.6). Como decisión metodológica propia, separa la estimación de las direcciones canónicas, con la calibración, de la selección y la proyección, que se repiten en el segmento que se limpia.

El EEG y $\boldsymbol{R}_{\text{lag}}$ se estandarizan con los estadísticos de la calibración (ecuación (3.4)), y con las matrices estandarizadas de calibración se estiman $n_{\text{comp}} \leq \min\{C, (2L+1)Q_{\text{EMG}}\}$ pares canónicos mediante un algoritmo iterativo que aproxima la solución de la ecuación (3.26) (Sección 4.4).

En el segmento que se limpia, el EEG y las referencias se estandarizan con los mismos estadísticos y se proyectan con los pesos calibrados, $\boldsymbol{U} = \boldsymbol{X}_z\boldsymbol{\Omega}_X$ y $\boldsymbol{V} = \boldsymbol{R}_{\text{lag},z}\boldsymbol{\Omega}_R$, y para cada par se calcula su correlación en ese segmento, $\rho_i = \text{corr}(\boldsymbol{U}_{:,i}, \boldsymbol{V}_{:,i})$. Como los pesos se estimaron en la calibración, esta correlación no está ordenada y puede cambiar de signo; por eso, se seleccionan los pares cuya correlación, en valor absoluto, alcanza el umbral τ:

$$\mathcal{B} = \{i : |\rho_i| \geq \tau\}, \qquad \text{y si } \mathcal{B} = \varnothing, \quad \mathcal{B} = \left\{\arg\max_i |\rho_i|\right\}. \tag{3.28}$$

Si ningún par alcanza el umbral, se elimina el de mayor $|\rho_i|$, para que una selección vacía no deje la salida idéntica a la entrada. Como $|\rho_i| \geq \tau$ equivale a $\rho_i^2 \geq \tau^2$, un umbral τ corresponde a $\theta = \tau^2$ en la regla de iCanClean (Sección 3.5); las dos reglas difieren en que aquí las correlaciones se calculan con pesos calibrados y en que siempre se elimina al menos un par.

La actividad seleccionada del lado de las referencias, $\boldsymbol{V}_{:,\mathcal{B}}$, se proyecta sobre el EEG estandarizado mediante una regresión ridge:

$$\boldsymbol{P}_\alpha = \left(\boldsymbol{V}_{:,\mathcal{B}}^{\mathsf{T}}\boldsymbol{V}_{:,\mathcal{B}} + \alpha\boldsymbol{I}\right)^{-1}\boldsymbol{V}_{:,\mathcal{B}}^{\mathsf{T}}\boldsymbol{X}_z, \qquad \widehat{\boldsymbol{A}}_z = \boldsymbol{V}_{:,\mathcal{B}}\,\boldsymbol{P}_\alpha, \tag{3.29}$$

donde α es el coeficiente de regularización y $\widehat{\boldsymbol{A}}_z$ es la parte del EEG estandarizado que se predice linealmente a partir de las variables seleccionadas. Luego, para cada canal c se calcula una ganancia

$$g_c = \min\left\{\max\left\{\frac{\langle \boldsymbol{x}_{z,c}, \widehat{\boldsymbol{a}}_{z,c}\rangle}{\|\widehat{\boldsymbol{a}}_{z,c}\|^2}, 0\right\}, g_{\max}\right\}, \tag{3.30}$$

donde $\boldsymbol{x}_{z,c}$ y $\widehat{\boldsymbol{a}}_{z,c}$ son las columnas c de $\boldsymbol{X}_z$ y $\widehat{\boldsymbol{A}}_z$. La ganancia acota la magnitud de la componente restada; como $\boldsymbol{P}_\alpha$ se ajusta en el mismo segmento, el cociente es al menos 1, y la ganancia solo reduce la resta cuando $g_{\max} < 1$.

La señal corregida en unidades originales es

$$\boldsymbol{X}_{\text{clean}} = \boldsymbol{X} - \gamma\,\widehat{\boldsymbol{A}}, \qquad \widehat{\boldsymbol{A}} = \widehat{\boldsymbol{A}}_z\,\text{diag}(\boldsymbol{g})\,\text{diag}(\boldsymbol{\sigma}_{\text{cal}}), \tag{3.31}$$

donde $\boldsymbol{g}$ reúne las ganancias, $\boldsymbol{\sigma}_{\mathrm{cal}}$ las desviaciones estándar de calibración de los canales EEG y γ es la escala de sustracción. Finalmente, si la limpieza aumenta la RMS de un canal en más de una fracción $\Delta_{\mathrm{máx}}$, el canal se reemplaza por la entrada sin modificar:

$$\boldsymbol{x}_{\mathrm{clean},c} \leftarrow \boldsymbol{x}_c \quad \text{si} \quad \mathrm{RMS}(\boldsymbol{x}_{\mathrm{clean},c}) > (1+\Delta_{\mathrm{máx}})\,\mathrm{RMS}(\boldsymbol{x}_c), \tag{3.32}$$

donde $\boldsymbol{x}_c$ y $\boldsymbol{x}_{\mathrm{clean},c}$ son las columnas c de $\boldsymbol{X}$ y $\boldsymbol{X}_{\mathrm{clean}}$, y la RMS se calcula sin centrar sobre todo el segmento. Con $\gamma \leq 1$, la resta no aumenta la suma de cuadrados de la señal estandarizada, de modo que la reversión solo actúa ante un cambio en la media del canal, como salvaguarda. La Figura 3.4 resume el procedimiento.

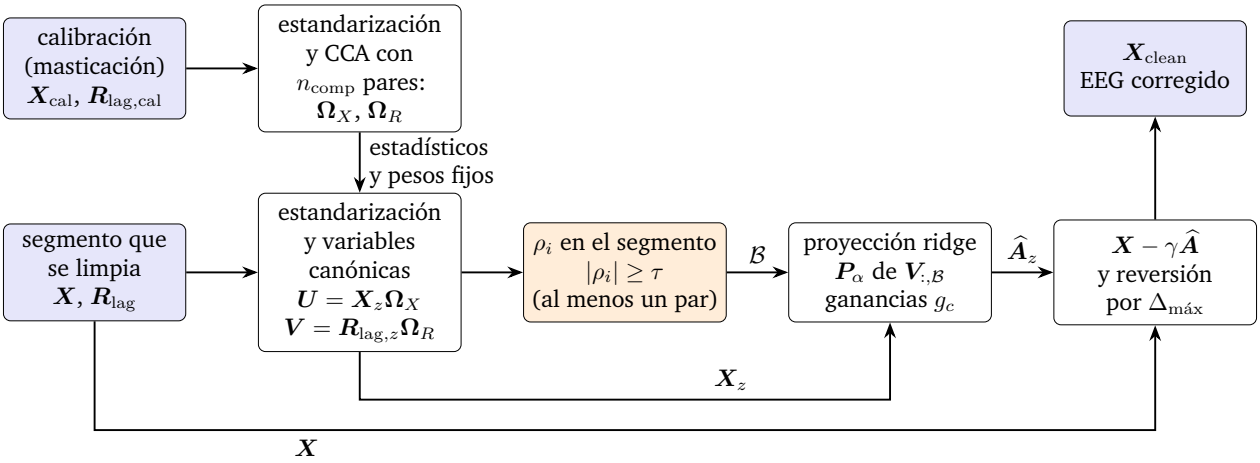


Figura 3.4: CCA con rezagos. Con la calibración se fijan la estandarización y los pesos canónicos $\boldsymbol{\Omega}_X$ y $\boldsymbol{\Omega}_R$. En el segmento que se limpia se calculan las correlaciones ρ_i, se seleccionan los pares con $|\rho_i| \geq \tau$ y se proyectan sus variables $\boldsymbol{V}_{:,\mathcal{B}}$ sobre el EEG estandarizado $\boldsymbol{X}_z$. La estimación se devuelve a las unidades del EEG y se resta de $\boldsymbol{X}$ con escala γ; un canal cuya RMS aumenta más de $\Delta_{\mathrm{máx}}$ conserva la entrada.

3.4.2. Limitaciones

CCA identifica covariación, no causalidad. Por lo tanto, una alta correlación EEG–EMG no demuestra que la componente EEG corresponda exclusivamente a actividad muscular. Una actividad neuronal que coincida temporalmente con la referencia puede ser seleccionada, al igual que un artefacto débil puede quedar sin corregir si no alcanza el umbral. Además, los pesos canónicos se estiman con una calibración breve, cuyas muestras consecutivas están muy correlacionadas, y con muchas columnas: los C canales EEG y hasta $(2L+1)Q_{\mathrm{EMG}}$ referencias rezagadas. Por eso, las correlaciones de la calibración pueden sobrestimar la relación entre ambos conjuntos, y las direcciones calibradas pueden ajustar actividad no relacionada con el artefacto o dejar de representarlo si la relación espacial entre EEG y referencias cambia en el tiempo. Por último, la componente forzada

hace que el método siempre reste una parte de la señal, incluso cuando ninguna correlación alcanza el umbral.

3.5. iCanClean

3.5.1. Definición

iCanClean utiliza CCA para identificar subespacios compartidos entre un registro contaminado y referencias de ruido adquiridas simultáneamente [11, 12]. El algoritmo consta de cuatro etapas: cálculo de las variables canónicas, selección de componentes, proyección de las componentes seleccionadas sobre los canales y sustracción directa de esa proyección. Estas etapas pueden aplicarse a una ventana larga o a ventanas cortas sucesivas, según el ruido sea estable o transitorio [11]. En esta tesis, el nombre iCanClean designa exclusivamente este algoritmo publicado.

En cada ventana, el EEG y las referencias se centran por columna con las medias de esa ventana, lo que da $\boldsymbol{X}_{\mathrm{MC}}$ y $\boldsymbol{R}_{\mathrm{MC}}$, y la CCA entre ambas matrices entrega los pesos $\boldsymbol{\Omega}_X$ y $\boldsymbol{\Omega}_R$ y las variables canónicas $\boldsymbol{U}$ y $\boldsymbol{V}$ de la ecuación (3.27), ordenadas por su correlación ρ_i. El número de pares está limitado por los rangos de las matrices centradas:

$$n_{\mathrm{par}} = \mín\left\{\mathrm{rank}(\boldsymbol{X}_{\mathrm{MC}}), \mathrm{rank}(\boldsymbol{R}_{\mathrm{MC}})\right\} \leq \mín\{C, Q_{\mathrm{EMG}}\}. \tag{3.33}$$

Por lo tanto, no supera el número de referencias, Q_{EMG}, y es menor si las referencias son linealmente dependientes [11].

3.5.2. Selección, proyección y sustracción

El algoritmo marca para eliminación los pares cuya correlación al cuadrado alcanza un umbral θ [11, 12]:

$$\mathcal{B} = \Big\{ i : \rho_i^2 \geq \theta \Big\}, \qquad 0 \leq \theta \leq 1. \tag{3.34}$$

La regla no fuerza la eliminación de un par: si $\mathcal{B}$ es vacío, la ventana no se modifica. La actividad seleccionada puede tomarse de las variables canónicas del EEG o de las de las referencias [11, 12]:

$$\boldsymbol{Z} = \boldsymbol{U}_{:,\mathcal{B}} \ \text{(fuente } \boldsymbol{U}\text{)} \qquad \text{o} \qquad \boldsymbol{Z} = \boldsymbol{V}_{:,\mathcal{B}} \ \text{(fuente } \boldsymbol{V}\text{)}. \tag{3.35}$$

Su proyección sobre los canales se estima por mínimos cuadrados ordinarios y se resta directamente:

$$\boldsymbol{P} = \arg\mín_{\boldsymbol{P}} \|\boldsymbol{X}_{\mathrm{MC}} - \boldsymbol{Z}\boldsymbol{P}\|_F^2 = \boldsymbol{Z}^{+}\boldsymbol{X}_{\mathrm{MC}}, \qquad \widehat{\boldsymbol{A}} = \boldsymbol{Z}\boldsymbol{P}, \qquad \boldsymbol{X}_{\mathrm{clean}} = \boldsymbol{X} - \widehat{\boldsymbol{A}}. \tag{3.36}$$

Como en la formulación publicada, la proyección no se regulariza y se resta completa [11]. En ambas fuentes, $\widehat{\boldsymbol{A}}$ es la proyección ortogonal de cada canal centrado sobre el subespacio generado por $\boldsymbol{Z}$, de modo que la energía total de cada canal centrado en la ventana no aumenta. Esta propiedad no

se extiende a cada banda de frecuencia ni a cada tramo de la ventana: la resta puede aumentar la potencia de una banda o la energía de un tramo más breve.

Las dos fuentes, sin embargo, no retiran lo mismo. Con la fuente $\boldsymbol{U}$, $\boldsymbol{Z} = \boldsymbol{X}_{\mathrm{MC}}\boldsymbol{\Omega}_{X,:,\mathcal{B}}$ es una combinación de los propios canales EEG: $\widehat{\boldsymbol{A}} = \boldsymbol{X}_{\mathrm{MC}}\boldsymbol{H}$, con $\boldsymbol{H} = \boldsymbol{\Omega}_{X,:,\mathcal{B}}\boldsymbol{P}$ de dimensión $C \times C$, y la limpieza equivale a aplicar a la ventana centrada el filtro espacial $\boldsymbol{I} - \boldsymbol{H}$, que elimina por completo, dentro de la ventana, la actividad de las $|\mathcal{B}|$ combinaciones espaciales seleccionadas, en todas las frecuencias, incluida la actividad neuronal que forme parte de ellas.

Con la fuente $\boldsymbol{V}$, en cambio, $\boldsymbol{Z} = \boldsymbol{R}_{\mathrm{MC}}\boldsymbol{\Omega}_{R,:,\mathcal{B}}$ es una combinación de las referencias, y $\widehat{\boldsymbol{A}}$ es la parte de cada canal que se predice linealmente a partir de ellas. La actividad del EEG que no está correlacionada con esas variables se conserva, aunque comparta su distribución espacial, y la forma de onda retirada hereda el contenido espectral de las referencias. Cuando ρ_i se aproxima a 1, $\boldsymbol{U}_{:,i}$ y $\boldsymbol{V}_{:,i}$ son casi proporcionales y ambas fuentes coinciden; con correlaciones moderadas, difieren.

Los autores sugieren en general la fuente $\boldsymbol{U}$, porque el artefacto se representaría mejor como un subespacio de los canales contaminados [12]. La evaluación considera ambas fuentes, junto con la duración de la ventana y el umbral θ (Sección 4.4). La Figura 3.5 resume el procedimiento.

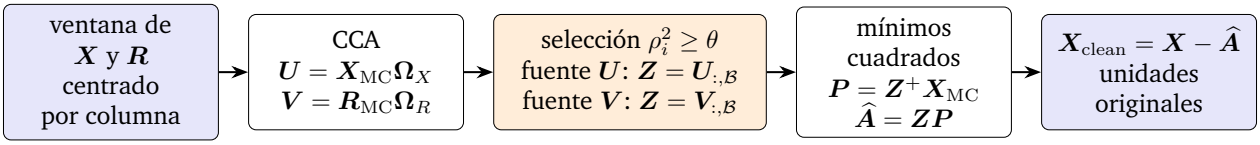


Figura 3.5: Etapas de iCanClean en cada ventana: centrado, CCA entre el EEG y las referencias, selección de los pares con $\rho_i^2 \geq \theta$, proyección por mínimos cuadrados de la actividad seleccionada y sustracción. La actividad seleccionada se toma de $\boldsymbol{U}$ o de $\boldsymbol{V}$. Si ningún par alcanza el umbral, la ventana no se modifica.

3.5.3. Limitaciones

La duración de la ventana introduce un compromiso: una ventana corta permite seguir cambios rápidos de la relación entre EEG y referencias, pero estima la CCA con menos muestras, de modo que las correlaciones calculadas varían más y pueden alcanzar el umbral por azar. La ventana y el umbral adecuados dependen, además, del tipo y del número de referencias. Si estas comparten el electrodo de referencia del EEG, la CCA puede seleccionar la actividad cerebral común a ambos conjuntos (Sección 2.5). Finalmente, como en la CCA con rezagos, una correlación alta no basta para atribuir al músculo la actividad retirada, por lo que conviene examinar tanto la señal corregida como la contribución eliminada.

3.6. Variante extendida propia

3.6.1. Definición

Esta tesis desarrolla una variante extendida propia, basada en una regresión ridge multisalida del EEG sobre referencias con rezagos y, opcionalmente, sobre su envolvente. A diferencia de CCA, aquí no se forman variables canónicas ni se seleccionan componentes por correlación canónica. Es un método distinto de iCanClean, con el que comparte el uso de referencias EMG para estimar y restar el artefacto. Tiene antecedentes en la regresión con canales auxiliares (Sección 2.4), de la que se distingue por usar rezagos de las referencias y regularizar los coeficientes; del filtrado adaptativo se distingue, además, por usar adelantos y mantener fijos los coeficientes calibrados.

3.6.2. Estimación regularizada

Sea $\tilde{\boldsymbol{R}}_{\mathrm{lag}}$ la matriz de rezagos de la referencia ampliada de la ecuación (3.8). El EEG y $\tilde{\boldsymbol{R}}_{\mathrm{lag}}$ se estandarizan según la ecuación (3.4), con lo que se obtienen $\boldsymbol{X}_z$ y $\tilde{\boldsymbol{R}}_{\mathrm{lag},z}$, y los coeficientes se ajustan con la calibración resolviendo

$$\boldsymbol{B}_\lambda = \arg\min_{\boldsymbol{B}} \left\{ \frac{1}{T_{\mathrm{cal}}} \left\| \boldsymbol{X}_{z,\mathrm{cal}} - \tilde{\boldsymbol{R}}_{\mathrm{lag},z,\mathrm{cal}} \boldsymbol{B} \right\|_F^2 + \lambda \left\| \boldsymbol{B} \right\|_F^2 \right\}, \tag{3.37}$$

cuya solución es

$$\boldsymbol{B}_\lambda = \left(\tilde{\boldsymbol{R}}_{\mathrm{lag},z,\mathrm{cal}}^{\mathsf{T}} \tilde{\boldsymbol{R}}_{\mathrm{lag},z,\mathrm{cal}} + \lambda\, T_{\mathrm{cal}}\, \boldsymbol{I} \right)^{-1} \tilde{\boldsymbol{R}}_{\mathrm{lag},z,\mathrm{cal}}^{\mathsf{T}} \boldsymbol{X}_{z,\mathrm{cal}}. \tag{3.38}$$

La división por T_{cal} en la ecuación (3.37) hace que λ no dependa de la duración de la calibración. Cada columna de $\boldsymbol{B}_\lambda$ corresponde a la regresión ridge de un canal EEG sobre los mismos predictores.

En el segmento que se limpia, las referencias se estandarizan con los mismos estadísticos de calibración y la estimación del artefacto se devuelve a las unidades del EEG:

$$\widehat{\boldsymbol{A}} = \left(\tilde{\boldsymbol{R}}_{\mathrm{lag},z} \boldsymbol{B}_\lambda \right) \mathrm{diag}(\boldsymbol{\sigma}_{\mathrm{cal}}), \qquad \boldsymbol{X}_{\mathrm{clean}} = \boldsymbol{X} - \gamma\, \widehat{\boldsymbol{A}}, \tag{3.39}$$

donde $\boldsymbol{\sigma}_{\mathrm{cal}}$ contiene las desviaciones estándar de calibración de los canales EEG y $\gamma \in (0, 1]$ es la escala de sustracción. La regularización λ controla la estabilidad del ajuste y γ, la intensidad de la sustracción. La Figura 3.6 resume el procedimiento.

La relación con iCanClean es más directa en un caso particular: con la fuente $\boldsymbol{V}$ y $\theta = 0$, iCanClean selecciona todos los pares, y la ecuación (3.36) se reduce a la regresión por mínimos cuadrados del EEG centrado sobre las referencias centradas de cada ventana. La variante parte de esa regresión, pero agrega rezagos y, opcionalmente, la envolvente; no selecciona por umbral; regulariza con λ y escala la sustracción con γ; y estima los coeficientes una sola vez, con la calibración. Estas diferencias no implican por sí mismas una mejor limpieza.

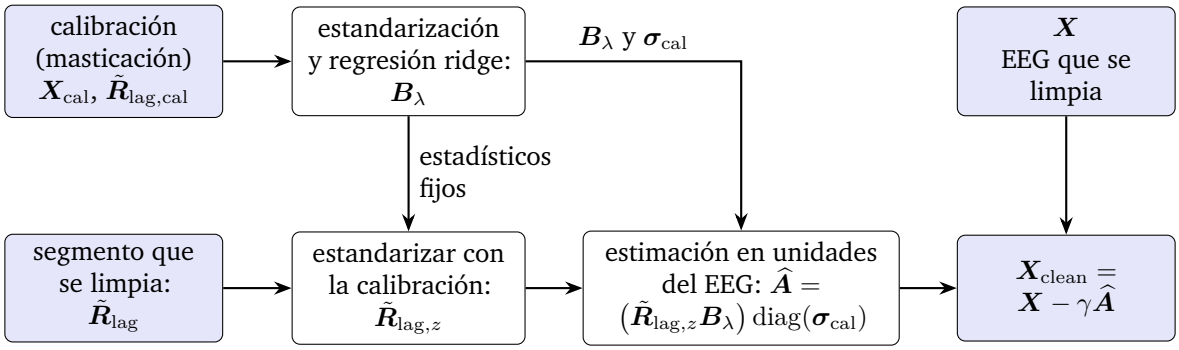


Figura 3.6: Variante extendida propia. Con la calibración se fijan la estandarización y los coeficientes ridge $\boldsymbol{B}_\lambda$. En el segmento que se limpia, las referencias con rezagos entre $-L$ y L se estandarizan con esos estadísticos y se multiplican por $\boldsymbol{B}_\lambda$; la estimación se devuelve a las unidades del EEG y se resta del EEG observado con escala γ. Según la configuración, $\tilde{\boldsymbol{R}}$ incluye la envolvente, calculada a partir del EMG filtrado en 20–120 Hz.

3.6.3. Limitaciones

La regresión supone que la contribución miogénica puede aproximarse mediante una combinación lineal de las referencias EMG con desfases temporales. Si los canales de referencia contienen actividad cerebral o no capturan la fuente miogénica dominante, la estimación resulta deficiente. El parámetro λ regula el compromiso entre el ajuste a los datos y la magnitud de los coeficientes: valores muy pequeños aumentan la sensibilidad del ajuste a la colinealidad entre los rezagos, y valores mayores pueden reducir la contribución estimada del artefacto. Los rezagos amplían la capacidad temporal del modelo, pero aumentan el número de predictores, $(2L+1)\tilde{Q}$, y con él el riesgo de ajustar en la calibración actividad no relacionada con el artefacto. Finalmente, como los coeficientes se fijan en la calibración, la variante supone que la relación entre referencias y EEG se mantiene durante la evaluación.

3.7. Relación entre los métodos y criterios de comparación

Los métodos estiman o atenúan la contaminación desde perspectivas diferentes. ASR utiliza una referencia estadística del EEG de reposo y reconstruye las direcciones que se desvían de ella. AF utiliza las referencias y actualiza en el tiempo un filtro lineal para cada canal. La CCA con rezagos proyecta sobre los canales las combinaciones de máxima correlación con las referencias rezagadas que selecciona. iCanClean aplica el mismo principio canónico, pero estima todas sus etapas en cada ventana. La variante extendida ajusta directamente una regresión ridge con referencias rezagadas y, opcionalmente, con su envolvente. Estas diferencias determinan qué parámetros y qué fallos son relevantes en cada caso. La Tabla 3.1 resume qué estima cada uno con la calibración y en el segmento que limpia.

Método	**Referencias EMG**	**Calibración**	**Segmento que se limpia**
ASR	No	$\boldsymbol{M}$ y $\boldsymbol{\Theta}$, con reposo	$\boldsymbol{\Sigma}_w$, $\delta_{w,j}$ y $\boldsymbol{G}_w$ en cada ventana
AF por bloques	$\tilde{\boldsymbol{R}}$ con rezagos causales	estandarización y pesos iniciales, con masticación	actualización de $\boldsymbol{W}_b$ y centrado de la predicción
CCA con rezagos	$\boldsymbol{R}_{\mathrm{lag}}$	estandarización, $\boldsymbol{\Omega}_X$ y $\boldsymbol{\Omega}_R$, con masticación	ρ_i, $\mathcal{B}$, $\boldsymbol{P}_\alpha$, g_c y reversión
iCanClean	$\boldsymbol{R}$	no se calibra	centrado, CCA, $\mathcal{B}$ y $\boldsymbol{P}$ en cada ventana
Variante extendida	$\tilde{\boldsymbol{R}}_{\mathrm{lag}}$	estandarización y $\boldsymbol{B}_\lambda$, con masticación	solo la aplicación de $\boldsymbol{B}_\lambda$

Tabla 3.1: Resumen de los métodos comparados: uso de referencias, magnitudes estimadas con la calibración y magnitudes estimadas en el segmento que se limpia.

La comparación debe conservar una lectura conjunta. Una reducción elevada de RMS puede acompañarse de una pérdida de actividad cerebral, mientras que una similitud elevada con el EEG observado puede dejar artefactos residuales. Por ello, se informan simultáneamente medidas de similitud y de reducción, separando los segmentos de reposo y masticación. La interpretación de un método se apoya en la relación entre esas medidas, en el espectro y en el supuesto operacional que las produce, sin convertir un único indicador en una clasificación de superioridad. La prueba semisintética complementa esta lectura con una señal objetivo conocida (Sección 4.7).

CAPÍTULO

4

Marco metodológico

En este capítulo se detalla la metodología experimental aplicada para evaluar y comparar el desempeño de métodos de remoción de artefactos electromiográficos de masticación en señales EEG. Se describen primero el conjunto de datos, el preprocesamiento común y la división de los registros en segmentos de calibración y de evaluación, y luego la implementación y los parámetros de cada método. Finalmente, se definen las métricas, que describen el compromiso entre la reducción de amplitud en masticación y la modificación del reposo, y el procedimiento de comparación. Estas decisiones conforman el protocolo de evaluación, cuya identificación se presenta en el Apéndice E.

4.1. Descripción del conjunto de datos

Los registros utilizados en esta tesis corresponden a los registrados por Sebastián Espinoza et al. Estos consisten en 15 registros en formato EEGLAB, cada uno con un archivo `.set` y uno `.fdt`, de seis participantes: tres del Experimento 1 (E1) y tres del Experimento 2 (E2). La comparación incluye cinco participantes: E1M1_R2, E1M2_R2, E2E3S7, E2E3S8 y E2E3S9 (alias E2s3s9). E1M3_R1 no se incluye, porque al anotar los intervalos (Sección 4.1.5) no se identificó en su registro un intervalo de masticación confiable, y su exclusión no depende del desempeño de los métodos. Aparte de la comparación, en ese registro se buscan posibles intervalos de masticación con el EMG de los maseteros, y los candidatos se informan como tales, sin modificar el conjunto de participantes.

4.1.1. Adquisición

La descripción del registro proviene de la información del laboratorio y de la documentación del fabricante. Las señales se adquirieron con un sistema BioSemi de electrodos activos, que, en lugar del electrodo de tierra de los equipos convencionales, utiliza dos electrodos con funciones separadas:

el electrodo de modo común (CMS), activo, y el de pierna derecha impulsada (DRL), pasivo. Ambos forman un lazo de realimentación que acerca el potencial medio del sujeto, es decir, la tensión de modo común, al potencial de referencia del conversor analógico-digital [65].

Los 64 electrodos EEG se ubicaron según el sistema 10-20 (Sección 1.3.1), siguiendo las instrucciones del fabricante, con gel conductor en los portaelectrodos del gorro. Como estos electrodos toleran impedancias altas, no se requiere la abrasión de la piel, y el contacto de cada uno puede comprobarse de forma continua con los valores de *offset* del software de adquisición [69]. Para ubicar los electrodos EMG, se pidió al participante un apriete dentario máximo y se localizó el masetero por palpación.

4.1.2. Estructura general del conjunto de datos

Cada registro contiene 72 canales: 64 canales de EEG, 4 canales EOG, destinados a la actividad electrooculográfica, y 4 canales situados sobre el músculo masetero superficial, para la actividad electromiográfica (EMG). Los rótulos de los canales EEG siguen la nomenclatura del sistema 10-20 extendido, en el orden de la plantilla de 64 canales de BioSemi incluida en MNE-Python [1].

La frecuencia de muestreo, de 1024 Hz en todos los registros, se conserva durante todo el análisis y deja un margen amplio respecto del límite de Nyquist, de 512 Hz, para las bandas utilizadas: 1–45 Hz en el EEG y hasta 120 Hz en las referencias EMG. Las señales se analizan en microvoltios (μV), la unidad en que se almacenaron, y las métricas, que son razones o correlaciones, no dependen de la unidad.

4.1.3. Distribución por experimento

El conjunto de datos se divide en dos experimentos con estructuras distintas:

- **Experimento 1 (E1):** el protocolo de cada participante comprende tres condiciones: reposo (*resting state*), masticación y una tarea realizada sin masticación. Se utiliza el archivo `raw` de cada participante, que contiene en un mismo registro el reposo y la masticación y sobre el cual se anotaron los intervalos de la Sección 4.1.5. Los archivos `ch` y `rs`, de unos 8 min, no contienen esos intervalos y no se evalúan; el examen de los archivos `ch` con el EMG, fuera de la evaluación, se describe en el Apéndice G.

- **Experimento 2 (E2):** cada participante cuenta con el registro completo (CH), que contiene la masticación, y un archivo de reposo (RS), distintos de los archivos `ch` y `rs` de E1. La tarea se realiza de manera sincronizada con la masticación. El archivo RS es una copia exacta del tramo inicial del archivo CH, de modo que ambos comparten origen y no son grabaciones independientes. Los dos archivos se emparejan por participante y se procesan por separado, sin concatenarlos.

Esta diferencia permite observar los métodos en registros con distinta organización temporal: en E1, la tarea se realiza por separado de la masticación, mientras que en E2 ambas ocurren de forma acoplada.

4.1.4. Tipos de canales

Los canales del conjunto de datos se organizan en tres grupos funcionales, según el mapa de canales auxiliares de estos registros:

- **Canales EEG:** 64 canales destinados al registro de la actividad electroencefalográfica sobre el cuero cabelludo.
- **Canales EOG:** 4 canales auxiliares, EXG1–EXG4, destinados al registro de la actividad electrooculográfica asociada a los movimientos oculares. EXG1 y EXG2 se ubican sobre y bajo el ojo derecho, y EXG3 y EXG4, a la izquierda y a la derecha de los ojos.
- **Canales EMG:** 4 canales auxiliares monopolares, EXG5–EXG8, ubicados sobre el músculo masetero superficial: EXG5 y EXG6 sobre el masetero izquierdo, en posición superior e inferior, y EXG7 y EXG8 sobre el derecho, en las mismas posiciones.

La Figura 4.1 muestra de forma esquemática la disposición de los electrodos EOG y EMG, con los lados referidos al participante.

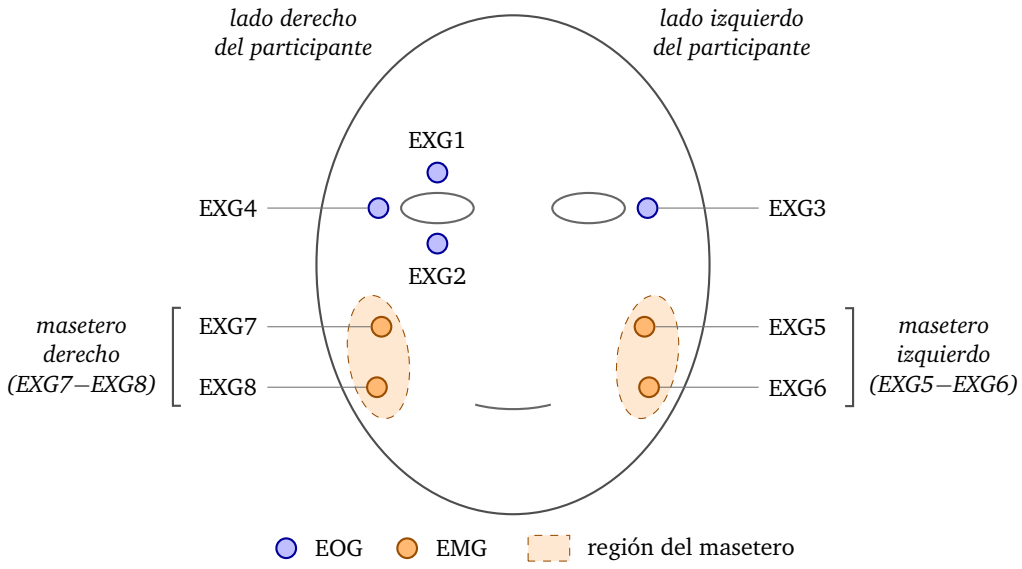


Figura 4.1: Disposición esquemática de los electrodos EOG (EXG1–EXG4) y EMG (EXG5–EXG8), en vista frontal: el lado derecho del participante aparece a la izquierda de la figura. Las referencias EMG principales son las derivaciones EXG5−EXG6 y EXG7−EXG8, y el control ocular utiliza EXG1−EXG2. Las posiciones son esquemáticas: siguen la descripción del montaje y no corresponden a la ubicación medida en cada participante. Elaboración propia.

Los canales EMG son fundamentales para esta tesis, porque proporcionan las referencias de artefacto que utilizan el AF por bloques, la CCA con rezagos, iCanClean y la variante extendida, mientras que ASR opera solo sobre la estructura estadística del propio EEG. Los canales EOG no se utilizan como referencias de limpieza ni intervienen en la selección de configuraciones, porque la comparación se centra en el artefacto muscular de masticación y agregar referencias oculares mezclaría la corrección de otro tipo de artefacto. Se utilizan solo en un control descriptivo (Sección 4.5.7).

4.1.5. Intervalos de reposo y masticación

Las condiciones se delimitan con intervalos anotados en segundos desde el inicio de cada archivo almacenado, como intervalos semiabiertos [inicio, fin). Los códigos numéricos de eventos de los archivos no se utilizan, porque no se dispone de un diccionario que los relacione con las condiciones. La Tabla 4.1 resume los 12 intervalos utilizados. Los dos intervalos de E1M1_R2 se anotaron manualmente, y los otros diez se propusieron con ayuda de una herramienta de inteligencia artificial, a partir de los cuatro canales EMG y del EEG (Apéndice H); su condición y su extensión se ratificaron visualmente.

Participante	Archivo	Masticación (s)	Reposo (s)
E1M1_R2	`E1M1_R2_raw_step0`	[799, 825)[a]	[826, 880)
E1M2_R2	`E1M2_R2_raw_step0`	[801, 837); [1045, 1082); [1294, 1352)	[900, 960)
E2E3S7	`E2E3S7_CH` y `E2E3S7_RS`	[1700, 1760) en CH	[30, 90) en RS
E2E3S8	`E2E3S8_CH` y `E2E3S8_RS`	[2000, 2060) en CH	[30, 90) en RS
E2E3S9[b]	`E2E3S9_CH` y `E2E3S9_RS`	[1680, 1740) en CH	[30, 90) en RS

Tabla 4.1: Intervalos de masticación y de reposo utilizados, en segundos desde el inicio de cada archivo. [a]Contiene un corte del registro en 815,91 s. [b]Alias: E2s3s9.

Los tramos de masticación se caracterizan por ráfagas repetidas de actividad en los cuatro canales EMG, y los de reposo, por una actividad EMG baja y sin aumentos sostenidos. Sin embargo, el reposo no equivale a un EEG libre de actividad muscular: la actividad de unidades motoras de los músculos frontal y temporal puede producir interferencia en los electrodos frontales y frontotemporales incluso en reposo [31].

4.2. Preprocesamiento de señales EEG

El preprocesamiento es una secuencia común, aplicada a todos los registros antes de los métodos de limpieza (Sección 3.1.2), cuyos pasos muestra la Figura 4.2. Los archivos se leen en Python, se comprueban la frecuencia de muestreo y el número y el orden de los canales, y los tipos de canal se asignan con el mapa de la Sección 4.1. Los 64 canales EEG se conservan, sin retirar ni interpolar

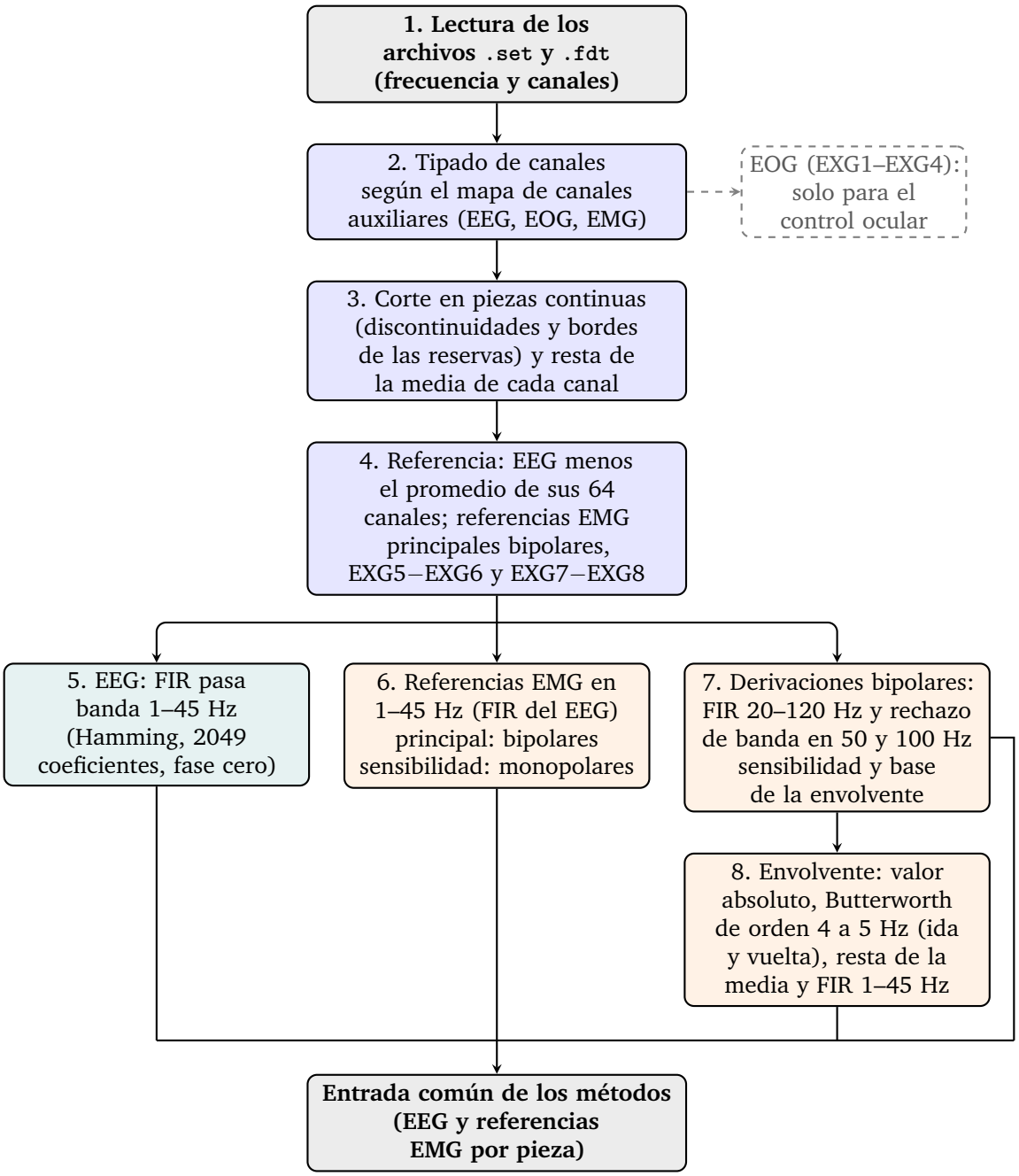


Figura 4.2: Secuencia de preprocesamiento común, aplicada a cada pieza por separado. La referencia principal, una derivación bipolar por masetero, comparte la banda del EEG; los canales monopolares (referidos al promedio del EEG o al electrodo CMS) y las derivaciones bipolares en 20–120 Hz, base también de la envolvente, se usan como sensibilidad. La cuarta alternativa, los canales monopolares almacenados en 20–120 Hz sin rechazos de banda, no se muestra. Elaboración propia.

canales, para que la comparación no dependa de una selección de canales favorable a algún método; el control de calidad exige, en cambio, valores finitos y canales no constantes.

Cada intervalo se divide en piezas continuas en las marcas de discontinuidad y en el borde de las reservas de calibración (Sección 4.3), y cada pieza se preprocesa por separado. Antes de filtrar, a cada canal se le resta su media, porque los desplazamientos de continua de las señales almacenadas son mucho mayores que la actividad registrada, y el filtro del EEG los atenúa en unos 48 dB, pero no los elimina. Sin el centrado, el residuo alteraría la entrada de métodos como ASR, que supone una señal de media nula [67, 68]. Después del filtrado se comprueba que, en cada canal, el valor absoluto de la media de la zona de evaluación no supere el 5 % de su desviación estándar; si lo supera, se informa, sin modificar el preprocesamiento.

Después del centrado se fija la referencia. El sistema BioSemi no aplica una referencia al guardar los datos: cada canal es la diferencia de potencial entre su electrodo y el electrodo CMS, y solo después de elegir una referencia se alcanza el rechazo de modo común completo del sistema [65]. Para el EEG, el laboratorio utilizó la referencia promedio, una de las habituales según el fabricante [65]; por ello, en cada muestra se resta a los 64 canales EEG su promedio, en el que no intervienen los canales EOG ni EMG. El EEG resultante tiene rango 63, lo que afecta a los métodos que operan sobre la covarianza entre canales, como ASR.

Como la referencia promedio reparte en todos los canales una parte de la actividad de cada uno, un canal de amplitud muy superior a la del resto modifica la entrada de todos los métodos. Por ello, en cada pieza se marca un canal si su RMS, calculada con el EEG centrado, referido al electrodo CMS, filtrado en 1–45 Hz y sin las guardas, supera 20 veces la mediana de las RMS de los 64 canales, un umbral convencional. Los canales marcados se conservan en el promedio, y su efecto se examina como análisis de sensibilidad.

El EEG se filtra con un FIR pasa banda de 1–45 Hz, diseñado por el método de ventana, con ventana de Hamming y 2049 coeficientes, es decir, de orden 2048, y frecuencias de corte a 6 dB de atenuación. El filtro se aplica como una convolución centrada y, como sus coeficientes son simétricos, tiene fase cero: no desplaza temporalmente los eventos registrados, una propiedad importante cuando se comparan los mismos eventos antes y después de la limpieza. En los extremos de cada pieza, la señal se extiende por reflexión con 1024 muestras. No se aplica un filtro notch, porque la interferencia de la red eléctrica queda fuera de la banda de paso, con una atenuación superior a 60 dB en 50 Hz; como comprobación, se calcula en cada canal la potencia relativa en 49–51 Hz después del filtro (Apéndice A.1).

Las referencias EMG principales son dos derivaciones bipolares, EXG5−EXG6 para el masetero izquierdo y EXG7−EXG8 para el derecho, de modo que $Q_{\mathrm{EMG}} = 2$. Como los canales EMG son monopolares y ambos canales de cada derivación están almacenados respecto del electrodo CMS, al restarlos ese potencial se cancela, y la derivación representa la diferencia de potencial entre dos puntos del mismo músculo. Así, las referencias principales no comparten la referencia del EEG, una propiedad relevante para los métodos basados en CCA (Sección 2.5).

La referencia principal se filtra en 1–45 Hz con el mismo filtro del EEG, porque, cuando el artefacto se estima como una combinación lineal de las referencias, como en el AF, en la CCA con rezagos, en la variante extendida y en la fuente $\boldsymbol{V}$ de iCanClean, la estimación hereda su contenido en frecuencia: unas referencias en 20–120 Hz solo permitirían representar el artefacto entre 20 y 45 Hz. Además, la correlación en que se basan la CCA e iCanClean depende de que ambas señales compartan contenido en frecuencia. Sin embargo, según la revisión de Muthukumaraswamy, que recoge las mediciones de O'Donnell et al., la actividad del masetero tiene un límite inferior cercano a 15 Hz y una frecuencia de pico en torno a 50–60 Hz [13], de modo que el contenido de 1–15 Hz de las referencias puede provenir también del movimiento o de actividad cerebral. El control con referencias de otro participante (Sección 4.5.7) permite examinar este riesgo.

Como análisis de sensibilidad se consideran cuatro alternativas a la referencia principal: los canales monopolares EXG5–EXG8 ($Q_{\mathrm{EMG}} = 4$) en 1–45 Hz, referidos al promedio del EEG, con lo que comparten su referencia, o tal como se almacenaron; las derivaciones bipolares en 20–120 Hz, una banda incluida en el ancho de banda típico de la actividad muscular, de aproximadamente 20–300 Hz [13], con un FIR del mismo diseño y los rechazos de banda de la envolvente; y los canales monopolares almacenados, en 20–120 Hz sin rechazos de banda («Monopolares, 20–120 Hz»). La comparación con los canales monopolares muestra el efecto de utilizar referencias que comparten actividad con el EEG.

La envolvente del EMG (ecuación (3.7)) se calcula en cada pieza a partir de las derivaciones principales, con un filtro Butterworth pasa-bajos de orden 4 y $f_{\mathrm{env}} = 5$ Hz y los demás pasos de la Figura 4.2. Los rechazos de banda en 50 y 100 Hz se aplican antes del valor absoluto, porque la rectificación, que es no lineal, convertiría las variaciones de amplitud de la interferencia de la red eléctrica en componentes de baja frecuencia.

4.3. Segmentos de calibración y evaluación

ASR se calibra con reposo, porque supone que la calibración representa la variabilidad de la señal que se desea conservar. El AF por bloques, la CCA con rezagos y la variante extendida estiman, en cambio, la relación entre las referencias EMG y el EEG; por eso se calibran con masticación, el segmento en que el artefacto es más visible. Esa calibración contiene también la actividad cerebral de la masticación y, en E2, la asociada a la tarea, y el modelo se ajusta con la misma condición en que después se mide la reducción del artefacto, aunque con muestras distintas. iCanClean y la identidad no se calibran.

Cada intervalo anotado se divide primero en piezas continuas en las marcas de discontinuidad del archivo; por ejemplo, el intervalo de masticación de E1M1_R2, [799, 825) s, contiene un corte en 815,91 s y forma dos piezas, que se procesan de forma independiente. Para calibrar ASR se reservan los primeros 34 s de la primera pieza de reposo que tenga al menos esa duración, con 30 s centrales de calibración y guardas de 2 s; para los otros métodos calibrados, los primeros 10 s de la primera

pieza de masticación que tenga al menos esa duración, con 6 s centrales. Cada reserva se separa del resto de su pieza y se procesa como una pieza propia.

El resto de esas piezas y las demás se destinan a la evaluación, también con guardas de 2 s, que superan el semisoporte del FIR y la ventana de ASR con su anticipación. Las guardas se filtran como contexto, pero no se utilizan para estimar parámetros ni se evalúan. La envolvente, en cambio, encadena filtros con un soporte combinado de unos 3,6 s, mayor que la guarda, por lo que puede presentar transitorios en la primera y en la última ventana evaluada de cada pieza. Este efecto no se corrige, porque ampliar las guardas o cambiar los filtros modificaría las ventanas evaluadas o la propia envolvente.

Sin las guardas, la zona de evaluación de cada pieza se divide en ventanas de 2 s sin solape, alineadas con su inicio, y el resto final, menor que una ventana, se descarta. Todas las configuraciones de todos los métodos se evalúan sobre exactamente las mismas ventanas. La Figura 4.3 resume esta partición.

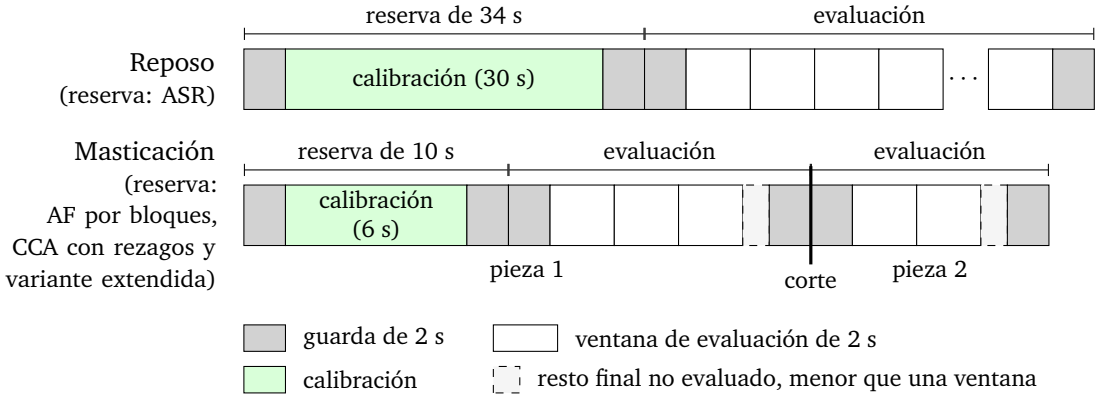


Figura 4.3: Reservas de calibración y ventanas de evaluación de un participante (esquema sin escala). Cada parte se filtra por separado, con guardas de 2 s que no se evalúan. La primera pieza elegible de cada condición aporta la reserva del método indicado entre paréntesis; el resto se divide en ventanas de evaluación de 2 s, comunes a todos los métodos. Elaboración propia.

Como la reserva forma una pieza propia, la calibración y la evaluación no comparten muestras ni se filtran juntas, y cuando provienen del mismo tramo continuo quedan separadas por 4 s, la suma de dos guardas. Los tramos reservados sirven también como conjunto de desarrollo: con ellos se aplica la regla sobre el término de sesgo del AF y se miden las propiedades de los métodos que se informan en el Apéndice A, de modo que esas mediciones no utilizan las ventanas de evaluación.

La calibración utiliza datos del propio participante. Así, todos los métodos calibrados siguen el régimen de ASR, cuya referencia estadística debe representar el reposo de la misma persona, y no se supone una relación entre las referencias EMG y el EEG común a varias personas, que estos datos no permiten verificar. Además, en el AF por bloques los pesos de la calibración son el punto de partida de la adaptación (Apéndice A.5), y la selección se simplifica, porque las calibraciones no dependen del participante que se deja fuera.

La separación temporal entre las condiciones difiere entre participantes. En E1M1_R2, el reposo comienza 1 s después del fin de la masticación. En E1M2_R2, el reposo, [900, 960) s, queda entre dos bloques de masticación: comienza 63 s después del primero y termina 85 s antes del segundo. En E2, el reposo es el tramo inicial del registro, en el archivo RS, y la masticación comienza entre 26,5 y 31,8 min después del fin del reposo.

4.4. Implementación y decisiones por método

Se comparan seis procedimientos: ASR, AF por bloques, CCA con rezagos, iCanClean, la variante extendida propia y la identidad, que devuelve la entrada sin cambios, $\boldsymbol{X}_{\text{clean}} = \boldsymbol{X}$, y sirve como línea base, porque cualquier modificación de la señal se mide respecto de ella. Esta sección describe la implementación y los parámetros de cada método, cuyo fundamento está en el Capítulo 3. Todos los métodos entregan los 64 canales EEG, y los métodos con referencia utilizan las dos referencias EMG principales. La Tabla 4.2 resume las configuraciones exploradas. Algunos parámetros del AF por bloques y de la variante extendida se expresan en muestras, con duraciones definidas para señales a 256 Hz; a 1024 Hz, esos valores se multiplican por cuatro, mientras que el paso y la fuga del AF, que actúan por actualización, se mantienen.

4.4.1. ASR

ASR se implementa con asrpy 0.0.8 [68], una implementación en Python de la versión euclidiana del método, y se calibra con la reserva de reposo de cada participante. La documentación de la biblioteca pide al menos 30 s de EEG razonablemente limpio y señala que el método suele utilizarse con 1 min o más [68]; la calibración de esta tesis se sitúa en ese mínimo. Como la biblioteca descarta las ventanas de calibración con amplitudes atípicas, la duración efectivamente utilizada se informa con los resultados (Tabla 5.1).

ASR recibe 63 de los 64 canales EEG, porque, con la referencia promedio, la matriz de covarianza de los 64 canales es singular, de rango 63, y la calibración de ASR requiere una matriz de rango completo. Por ello, dentro de ASR se retira el canal O2, que no es un canal primario y se elige sin examinar los datos, y después de la limpieza se reconstruye como menos la suma de los otros 63 canales de la salida. Como, con la referencia promedio, O2 coincide en cada muestra con esa combinación, retirarlo no elimina información, y O2 sigue presente en todas las salidas y en todas las métricas.

La calibración y cada pieza se dividen por una escala global, común a todos los canales, y la salida se multiplica de nuevo por ella; esa escala y dos adaptaciones de asrpy que no cambian su algoritmo se detallan en el Apéndice A.2. La ventana estadística es de 0,5 s, el valor predeterminado de la biblioteca, y la anticipación, de 0,25 s, la mitad de esa ventana, como recomienda su documentación [68]. La matriz de reconstrucción se actualiza cada 64 muestras, es decir, cada 62,5 ms, el doble del valor predeterminado de asrpy, de 32 muestras: un intervalo mayor reduce el

Método	Parámetros explorados	Parámetros fijos	Calibración	N.º
Identidad	—	—	—	1
ASR	$k \in \{5, 10, 15, 20, 25, 30\}$; $f_{\mathrm{ASR}} \in \{0{,}33;\ 0{,}66\}$	ventana de 0,5 s; anticipación de 0,25 s; actualización cada 64 muestras; 63 canales (sin O2, que se reconstruye)	30 s de reposo	12
AF por bloques	envolvente {sin, con}; $L_{\mathrm{AF}} \in \{64, 256\}$ muestras (16 y 64); $\mu \in \{0{,}01;\ 0{,}03;\ 0{,}1\}$	$\nu = 0{,}001$; $B = 8$ muestras (2); $\gamma = 1$; $f_{\mathrm{env}} = 5$ Hz; con término de sesgo	6 s de masticación	12
CCA con rezagos	$n_{\mathrm{comp}} \in \{2, 4, 6\}$; $L \in \{0, 1, 2\}$ muestras; $\tau \in \{0{,}10;\ 0{,}15;\ 0{,}20\}$	$\alpha = 0{,}01$; $\gamma = 1$; $g_{\mathrm{máx}} = 1$; $\Delta_{\mathrm{máx}} = 0{,}02$	6 s de masticación	27 (21)
iCanClean	ventana {2, 4} s; $\theta \in \{0{,}20;\ 0{,}35;\ 0{,}50;\ 0{,}65\}$; fuente $\boldsymbol{Z} \in \{\boldsymbol{U}, \boldsymbol{V}\}$	—	estimación por ventana	16
Variante extendida	envolvente {sin, con}; $\lambda \in \{0{,}001;\ 0{,}01;\ 0{,}1\}$; $\gamma \in \{0{,}6;\ 0{,}8;\ 1{,}0\}$; $L \in \{16, 32\}$ muestras (4 y 8)	$f_{\mathrm{env}} = 5$ Hz	6 s de masticación	36
Total				104

Tabla 4.2: Configuraciones exploradas por método. Cada combinación de los valores explorados de un método define una configuración, y los parámetros fijos conservan el mismo valor en todas ellas. Los valores en muestras corresponden a 1024 Hz; en el AF por bloques y en la variante extendida se indica entre paréntesis su equivalente a 256 Hz, mientras que el intervalo de actualización de ASR y los rezagos de la CCA con rezagos se fijan directamente a 1024 Hz. En la CCA con rezagos, $g_{\mathrm{máx}}$ es la ganancia máxima por canal y $\Delta_{\mathrm{máx}}$, el aumento relativo máximo de RMS, y las 27 configuraciones nominales corresponden a 21 distintas (Sección 4.4.3). En la variante extendida, λ regulariza el objetivo normalizado por el número de muestras de calibración (ecuación (3.37)); por lo tanto, equivale a una regularización $\lambda\, T_{\mathrm{cal}}$ en la formulación sin normalizar. La duración de calibración es la reservada, sin las guardas. iCanClean se calcula además con ventanas de 8 s, en ocho configuraciones que solo intervienen en un análisis de sensibilidad (Sección 4.5.10). La raya (—) indica que el elemento no corresponde al método.

costo de cómputo [68], y este se mantiene por debajo de los aproximadamente 100 ms entre actualizaciones descritos por Mullen et al. [46].

Se exploran la tolerancia k, que la biblioteca denomina `cutoff`, con valores menores y mayores que su valor predeterminado, $k = 20$, y la fracción dimensional f_{ASR}, denominada `maxdims`. En los datos de desarrollo, cambios mínimos de la calibración modificaron la salida de ASR (Apéndice A.2); por eso, con la configuración elegida, ASR se repite con las variaciones de la Sección 4.5.10.

4.4.2. AF por bloques

Las referencias, con o sin envolvente, y el EEG se estandarizan con los estadísticos de la reserva de masticación, con la que también se obtienen los pesos iniciales. Cada pieza de evaluación parte de una copia de esos pesos, con la historia de los rezagos en cero, y la adaptación continúa durante toda la pieza; el artefacto estimado se centra y se resta con escala $\gamma = 1$ (ecuación (3.25)). Se exploran la envolvente, la memoria L_{AF} y el paso μ. La fuga ν y el avance B son valores fijos de esta implementación, que no proceden de una fuente bibliográfica. La inclusión del término de sesgo se decide con una regla aplicada a los datos de desarrollo (Apéndice A.5), según la cual todas las configuraciones lo incluyen.

4.4.3. CCA con rezagos

Las correlaciones canónicas se calculan con la implementación de scikit-learn, que utiliza un algoritmo iterativo NIPALS. El modelo se ajusta con la reserva de masticación sobre el EEG y sobre las referencias rezagadas $\boldsymbol{R}_{\mathrm{lag}}$ (ecuación (3.6)), estandarizados con los estadísticos de la calibración. Se exploran el número de componentes, el rezago máximo L, de hasta 2 muestras, es decir, desfases de hasta unos 2 ms, y el umbral τ; con rezagos mayores, en los datos de desarrollo, la modificación quedó confinada a los bordes de cada pieza (Apéndice A.3). Como el número de componentes se limita al de columnas de referencia, con $L = 0$ los valores $n_{\mathrm{comp}} = 4$ y $n_{\mathrm{comp}} = 6$ equivalen a $n_{\mathrm{comp}} = 2$. Los valores de α, γ, $g_{\mathrm{máx}}$ y $\Delta_{\mathrm{máx}}$ son fijos de esta implementación y no proceden de una fuente bibliográfica. Como la regla de la ecuación (3.28) retira siempre al menos el par de mayor $|\rho_i|$, la selección no basta para atribuir al artefacto el par retirado, cuya especificidad se examina con el control (a) (Sección 4.5.7).

4.4.4. iCanClean

iCanClean se implementa según el algoritmo publicado por Downey y Ferris [11, 12]. En cada ventana se estiman de nuevo la CCA entre el EEG y las referencias, mediante la descomposición en valores singulares de las señales centradas, la selección de pares y la proyección, de modo que el método no requiere calibración. Con las dos referencias principales, cada ventana tiene a lo sumo dos pares canónicos (ecuación (3.33)), que se seleccionan con $\rho_i^2 \geq \theta$. Las diferencias con el ventaneo de la implementación oficial se detallan en el Apéndice A.4.

Las ventanas de iCanClean no se solapan y se alinean con el inicio de la zona de evaluación, de modo que cada ventana evaluada queda limpiada por una sola ventana de iCanClean; con ventanas de 4 s y un número impar de ventanas de evaluación, la última se extiende 2 s hacia la guarda final. Como en los valores predeterminados de la implementación oficial para la línea de comandos [60], cada ventana se limpia con la proyección estimada en ella, por lo que la salida puede presentar discontinuidades entre ventanas. Por ello, en cada borde y en cada canal primario se mide el salto de la estimación del artefacto, $|\widehat{a}_c[n] - \widehat{a}_c[n-1]|$, con n la primera muestra de la ventana siguiente,

en relación con la RMS de la entrada en las dos ventanas contiguas, y se informan su mediana y su máximo.

Se exploran la duración de la ventana, el umbral θ y la fuente de la actividad que se resta. Las duraciones de 2 y 4 s fueron las más prometedoras en el barrido de Gonsisko et al., en el que el umbral óptimo disminuyó de 0,65 con 120 canales de referencia a 0,25 con 16 [58]; como aquí se dispone de solo dos referencias principales, la grilla se extiende hacia valores menores. Con umbrales bajos y ventanas breves, las correlaciones pueden alcanzar el umbral por azar (Sección 3.5). Por ello, para cada duración y condición se informan la distribución de ρ_1^2 y la fracción de ventanas con alguna componente retirada, con las referencias propias y con las del control (a) (Sección 4.5.7).

4.4.5. Variante extendida propia

La variante extendida (Sección 3.6) se calibra con la reserva de masticación y se aplica sin cambios a la evaluación; su grilla está en la Tabla 4.2. No interviene en los criterios con que se examina la hipótesis (Sección 4.5.8).

4.5. Cálculo de métricas, selección y criterios de evaluación

Para comparar el desempeño de los métodos se utiliza un conjunto de métricas complementarias que describen dos aspectos, calculados por separado en cada condición: la reducción de amplitud durante la masticación y la modificación que cada método introduce en la señal durante el reposo. En este trabajo, se considera favorable el desempeño de un método cuando logra una reducción de amplitud durante la masticación acompañada, durante el reposo, de una alta similitud con el EEG observado y de una baja amplitud de la señal eliminada. Como no se dispone de una señal cerebral de referencia, las métricas describen cambios respecto del EEG observado y no miden directamente la preservación de la actividad cerebral. Por ello, se interpretan en conjunto y se complementan con la inspección visual de la Sección 4.6.

En las ecuaciones siguientes, para un canal c, $x[n] = (\boldsymbol{X})_{n,c}$ es la señal EEG preprocesada e $y[n] = (\boldsymbol{X}_{\text{clean}})_{n,c}$, la señal corregida por el método, con la notación de la Sección 3.1. El índice n recorre las N muestras de las ventanas de evaluación de una condición, y ambas señales se centran en cada ventana. Las métricas se calculan por canal, y la condición se indica con un subíndice, como en R^2_{reposo} o $\Delta\text{RMS}_{\text{mast}}$.

El análisis principal se resume en diez canales fijos y comunes a todos los métodos, los canales primarios: Fp1, Fpz, Fp2, AFz, AF3, AF4, AF7, Fz, T7 y T8. Este conjunto reúne derivaciones frontales y temporales, próximas a la musculatura facial y mandibular, y derivaciones de la línea media. Las métricas de los 64 canales se conservan para una descripción secundaria.

4.5.1. Similitud en reposo

El coeficiente de correlación de Pearson entre la entrada y la salida se calcula como

$$\mathrm{CC} = r_{xy} = \frac{\sum_n (x[n]-\bar{x})\,(y[n]-\bar{y})}{\sqrt{\sum_n (x[n]-\bar{x})^2 \, \sum_n (y[n]-\bar{y})^2}}, \tag{4.1}$$

donde $\bar{x}$ e $\bar{y}$ son las medias de cada señal. El coeficiente toma valores en $[-1,\,1]$, y un valor negativo indica una asociación invertida. A partir de él se define

$$R^2 = 100\, r_{xy}^2, \tag{4.2}$$

donde el factor 100 expresa el resultado como porcentaje. R^2 es el cuadrado de una correlación y no el coeficiente de determinación de un modelo ajustado. Como tal, es invariante a la escala y al signo: una salida $y[n] = -0{,}1\,x[n]$ obtiene $R^2 = 100$, aunque su RMS sea el 10 % de la entrada y su signo esté invertido. Por ello, R^2 se interpreta junto con la correlación con signo y con la razón de RMS. Si la salida es constante, la correlación es indefinida y R^2 se informa como un valor faltante.

4.5.2. Reducción de RMS durante masticación

Para una señal de N muestras, la raíz cuadrática media (RMS) cuantifica su amplitud eficaz. La razón entre la RMS de la salida y la de la entrada es

$$\mathrm{RMS}(x) = \sqrt{\frac{1}{N}\sum_{n=1}^{N} x[n]^2}, \qquad Q_{\mathrm{RMS}} = \frac{\mathrm{RMS}(y)}{\mathrm{RMS}(x)}. \tag{4.3}$$

$Q_{\mathrm{RMS}} = 1$ indica la misma amplitud eficaz; $Q_{\mathrm{RMS}} < 1$, atenuación; y $Q_{\mathrm{RMS}} > 1$, aumento. En reposo se informa $|1 - Q_{\mathrm{RMS,reposo}}|$ como medida del cambio de amplitud que el método introduce cuando la contaminación muscular es baja. La reducción porcentual de RMS es

$$\Delta\mathrm{RMS} = 100\big(1 - Q_{\mathrm{RMS}}\big). \tag{4.4}$$

Un valor positivo indica disminución de amplitud, y un valor negativo, que la RMS aumentó después de limpiar. Durante la masticación, $\Delta\mathrm{RMS}_{\mathrm{mast}}$ describe la reducción de amplitud, pero no permite establecer por sí sola que la componente eliminada corresponda exclusivamente al artefacto muscular.

4.5.3. Señal eliminada

La amplitud de la señal eliminada, $x[n] - y[n]$, en relación con la entrada se mide como

$$E_{\mathrm{rel}} = \frac{\|x - y\|}{\|x\|} = \frac{\mathrm{RMS}(x-y)}{\mathrm{RMS}(x)}. \tag{4.5}$$

Esta medida coincide con la raíz del error cuadrático medio relativo (RRMSE) respecto de la entrada, y su cuadrado es el error cuadrático medio normalizado (NMSE); por eso se informa solo E_{rel}. Aquí, el término error designa una diferencia respecto del EEG observado: una diferencia pequeña durante la masticación puede corresponder a un artefacto residual, y una grande puede incluir actividad cerebral además del artefacto.

4.5.4. Potencia por banda

La densidad espectral de potencia se estima con el método de Welch en cada ventana de evaluación de 2 s, y luego se promedian los espectros de las ventanas. El método utiliza segmentos de 1024 muestras, ventana de Hann, solapamiento de 512 muestras, transformada de 1024 puntos, eliminación de la media de cada segmento y densidad unilateral, con una resolución en frecuencia de 1 Hz. Si $\widehat{P}_x(f)$ y $\widehat{P}_y(f)$ son las densidades de la entrada y de la salida, la potencia en una banda $\mathcal{F} = [f_1,\, f_2]$ y su razón son

$$P_{\mathrm{in}}(\mathcal{F}) = \int_{\mathcal{F}} \widehat{P}_x(f)\, df, \qquad P_{\mathrm{out}}(\mathcal{F}) = \int_{\mathcal{F}} \widehat{P}_y(f)\, df, \qquad \frac{P_{\mathrm{out}}(\mathcal{F})}{P_{\mathrm{in}}(\mathcal{F})}, \tag{4.6}$$

donde la integral se calcula por la regla de los trapecios sobre los puntos de frecuencia comprendidos entre f_1 y f_2, ambos incluidos. Se consideran cinco bandas: δ (1–4 Hz), θ (4–8 Hz), α (8–13 Hz), β (13–30 Hz) y γ (30–45 Hz), que comparten sus bordes y, en conjunto, cubren el intervalo de 1–45 Hz sin vacíos. La razón se calcula por canal y luego se promedia sobre los canales; también se expresa en decibeles como $10 \log_{10}(P_{\mathrm{out}}/P_{\mathrm{in}})$. Además, se calcula la reducción de potencia en 20–45 Hz durante la masticación, que se utiliza en la Sección 4.5.6.

4.5.5. Interpretación de las métricas

En la Tabla 4.3 se resumen los indicadores empleados para evaluar el desempeño de cada método.

4.5.6. Puntaje balanceado

Para seleccionar configuraciones se utiliza el puntaje balanceado propuesto en este trabajo, que combina con igual peso la similitud en reposo y la reducción de amplitud durante la masticación:

$$S_{\mathrm{bal}} = \frac{1}{2} R^2_{\mathrm{reposo}} + \frac{1}{2} \Delta\mathrm{RMS}_{\mathrm{mast}}. \tag{4.7}$$

Ambos términos se calculan por participante, como media de los diez canales primarios. La identidad obtiene $R^2_{\mathrm{reposo}} = 100$ y $\Delta\mathrm{RMS}_{\mathrm{mast}} = 0$, es decir, $S_{\mathrm{bal}} = 50$, y un método que modifica la señal supera ese valor solo si la reducción de amplitud en masticación supera la pérdida de similitud en reposo. Sin embargo, el puntaje también premia una atenuación uniforme. Si la salida solo multiplica la entrada por una constante positiva menor que uno, su razón de RMS es esa misma constante en ambas condiciones y, como R^2 es invariante a la escala, se obtiene $R^2_{\mathrm{reposo}} = 100$ y

Métrica	Símbolo	Condición principal	Lectura
Correlación de Pearson	CC	ambas	Asociación lineal con signo entre entrada y salida.
Correlación al cuadrado	R^2	reposo	Alta: mayor similitud lineal con la entrada. No informa escala ni signo.
Razón de RMS	Q_{RMS}	reposo	Uno: igual amplitud eficaz; menor que uno: atenuación; mayor que uno: aumento.
Reducción de RMS	$\Delta\mathrm{RMS}$	masticación	Alta: mayor reducción de amplitud.
Amplitud relativa eliminada	E_{rel}	ambas	Alta: mayor diferencia respecto del EEG observado.
Razón de potencia por banda	$P_{\mathrm{out}}/P_{\mathrm{in}}$	ambas	Uno: sin cambio en la banda; menor que uno: atenuación.

Tabla 4.3: Métricas calculadas sobre las mismas ventanas de evaluación para todos los métodos.

$S_{\mathrm{bal}} = 50 + 50\,(1 - Q_{\mathrm{RMS}})$, sin que el artefacto se haya limpiado de forma selectiva. Por eso, S_{bal} se informa siempre junto con $Q_{\mathrm{RMS,reposo}}$, que sí refleja ese cambio de amplitud, y si un método supera a la identidad en S_{bal} se informa sin asignarle valor de mérito.

Como análisis de sensibilidad se calcula además

$$S_{\mathrm{bal,banda}} = \frac{1}{2} R^2_{\mathrm{reposo}} + \frac{1}{2} \cdot 100 \left(1 - \frac{\overline{P_{\mathrm{out}}(20\text{–}45\ \mathrm{Hz})}}{P_{\mathrm{in}}(20\text{–}45\ \mathrm{Hz})} \right)_{\mathrm{mast}}, \tag{4.8}$$

donde la barra indica la media de los diez canales primarios. Este puntaje reemplaza la reducción de RMS, que puede quedar dominada por las componentes lentas de mayor amplitud, por la reducción de potencia en 20–45 Hz, una banda en que un estudio con dos sujetos encontró actividad mayoritariamente muscular (Sección 1.4). Como una razón de potencia es el cuadrado de una razón de amplitud, ambos puntajes no están en la misma escala; $S_{\mathrm{bal,banda}}$ se calcula con la configuración elegida por S_{bal} y no interviene en la selección.

4.5.7. Selección de configuraciones y controles

Para cada método se elige una configuración por participante, de modo que el desempeño que se informa no provenga de los datos con que se eligió la configuración. Sea $\mathcal{P}$ el conjunto de los cinco participantes y $\mathcal{C}$ el de configuraciones de un método (Tabla 4.2). Para cada participante p se elige

$$\zeta^*(p) = \arg\max_{\zeta \in \mathcal{C}} \; \frac{1}{4} \sum_{p' \in \mathcal{P} \setminus \{p\}} S_{\mathrm{bal}}(p', \zeta), \tag{4.9}$$

donde $S_{\mathrm{bal}}(p', \zeta)$ es el puntaje balanceado de la ecuación (4.7) en el participante p' con la configuración ζ; es decir, la configuración con mayor puntaje medio en los otros cuatro participantes. Así, ningún parámetro calibrado ni ningún puntaje de selección proviene de las ventanas de evaluación de p, aunque varios métodos estiman parte de su modelo sobre la señal que limpian, como parte de su algoritmo y sin utilizar sus métricas.

Las configuraciones cuyo puntaje medio queda a menos de 0,5 puntos del máximo se consideran empatadas, y entre ellas se elige la que menos modifica el reposo, es decir, la de menor media de $|1 - Q_{\mathrm{RMS,reposo}}|$ en $\mathcal{P} \setminus \{p\}$. Si el empate persiste, decide un orden fijo: los parámetros se recorren en el orden de la Tabla 4.2, cada uno de menor a mayor, con «sin» antes que «con» y $\boldsymbol{U}$ antes que $\boldsymbol{V}$. Una configuración con un valor indefinido en algún participante de $\mathcal{P} \setminus \{p\}$ no es elegible, salvo que todas lo estén por una causa común, como una calibración de ASR que falla, caso en que ese participante se excluye del promedio de selección; si ninguna es elegible, el método queda sin configuración para p.

Para examinar si la limpieza depende de la información contenida en las referencias, la configuración elegida del AF por bloques, la CCA con rezagos, iCanClean y la variante extendida se aplica además con controles. En el control (a), el control principal de especificidad, las referencias de cada pieza se reemplazan por las de otro participante, de la misma condición (Apéndice A.6), y los métodos calibrados conservan su calibración propia: si la limpieza se mantiene, no puede atribuirse a la información de las referencias propias. Como este control aplica a las referencias ajenas una relación con el EEG estimada con las propias, que pueden tener otra amplitud, tiende a favorecer a las referencias reales. Por ello se agrega el control (a′), en que los métodos calibrados se calibran con el EEG de la reserva de masticación del participante y con las referencias de la reserva del otro; en iCanClean, que no se calibra, ambos controles coinciden.

En el control (b), las referencias se reemplazan por ceros, de modo que la salida debe ser idéntica a la entrada en iCanClean y en la variante extendida, y coincidir con ella salvo por el redondeo en la CCA con rezagos: en estos métodos, el control verifica la implementación. En el AF por bloques, el término de sesgo sigue actuando y el control muestra su efecto. Con $\mu = 0{,}1$, la atenuación en 1–8 Hz es comparable a la del sesgo solo; si la configuración elegida tiene ese paso, esa atenuación se lee junto con este control. Para cada control se informa también la fracción de la atenuación que se mantiene, $\Delta\mathrm{RMS}_{\mathrm{mast}}(\mathrm{a})/\Delta\mathrm{RMS}_{\mathrm{mast}}(\mathrm{propias})$, o su análoga con el control (a′), y, en iCanClean, el número de componentes retiradas por ventana.

Se aplica también un control descriptivo con el EOG, porque en la región frontal parte de la actividad de baja frecuencia registrada en masticación puede seguir a la actividad ocular. Para cada método y para la identidad se calculan en 1–8 Hz, como mediana de los siete canales frontales Fp1, Fpz, Fp2, AF7, AF3, AFz y AF4, la fracción de la varianza de la señal retirada, $x[n] - y[n]$, y la de la entrada que explica la derivación vertical EXG1−EXG2; la fracción de la varianza retirada que explican las referencias EMG principales; y la fracción retirada de la varianza de la entrada. El resultado se describe como actividad retirada compartida con el EOG, y no necesariamente como actividad ocular retirada. Como sensibilidad se utilizan la derivación horizontal EXG3−EXG4 y los

canales EOG monopolares, que en masticación comparten actividad con el EMG. Ninguno de estos controles interviene en la selección.

Además de las configuraciones elegidas, se informan sin selección la identidad (Tabla 5.7), tres configuraciones de referencia tomadas de la literatura y de los valores predeterminados del software (rótulos LIT) y cuatro configuraciones adicionales (rótulos HIS); su definición y sus métricas medias se presentan en el Apéndice B, salvo las de HIS-VAR (Tabla F.1).

4.5.8. Criterios para examinar la hipótesis

La hipótesis de esta tesis compara iCanClean con ASR, AF y CCA en la atenuación del artefacto durante la masticación y en la modificación de la señal en reposo. Aquí, la atenuación del artefacto se interpreta como una mayor reducción de la amplitud y de la potencia en 20–45 Hz del EEG registrado durante la masticación, y preservar la señal en reposo, como modificar menos el EEG observado en esa condición, aunque la reducción no indica por sí sola qué parte de lo retirado era artefacto. Con cinco participantes no es posible un contraste confirmatorio; por ello, la hipótesis se examina con criterios descriptivos. Estos criterios, al igual que las grillas y las demás decisiones de diseño de este capítulo, no son independientes de los registros evaluados: se definieron conociendo el comportamiento de los métodos, con otro preprocesamiento, en los mismos tramos temporales que 131 de las 178 ventanas evaluadas. Por eso, los criterios resumen la evidencia, pero no la confirman de forma independiente.

Con la configuración $\zeta^*(p)$ de cada método, los criterios son los siguientes:

1. **Atenuación en masticación.** En el mismo participante, $\Delta\mathrm{RMS}_{\mathrm{mast}}$ y la reducción de potencia en 20–45 Hz de iCanClean son mayores que las del comparador, al menos por el margen correspondiente.

2. **Modificación en reposo.** En el mismo participante, $100 - R^2_{\mathrm{reposo}}$, $|1 - Q_{\mathrm{RMS,reposo}}|$ y $\max_{\mathcal{F}} |10 \log_{10}(P_{\mathrm{out}}(\mathcal{F})/P_{\mathrm{in}}(\mathcal{F}))|$, calculado sobre las cinco bandas, son menores en iCanClean que en el comparador, al menos por el margen correspondiente.

3. **Requisito previo.** iCanClean supera a la identidad en $\Delta\mathrm{RMS}_{\mathrm{mast}}$, al menos por el margen correspondiente.

Los márgenes son de 1 punto porcentual para $\Delta\mathrm{RMS}_{\mathrm{mast}}$ y para la reducción de potencia en 20–45 Hz, de 1 punto para R^2_{reposo}, de 0,01 para $|1 - Q_{\mathrm{RMS,reposo}}|$ y de 0,5 dB para la razón de potencia por banda. Son márgenes convencionales, que no se derivaron de los datos; el de la reducción de potencia, que no está en la escala de $\Delta\mathrm{RMS}_{\mathrm{mast}}$, es comparativamente menos exigente. Cuando un criterio considera varias métricas, la comparación se cumple si todas sus diferencias alcanzan el margen a favor de iCanClean, no se cumple si alguna lo alcanza a favor del comparador y es un empate en los demás casos. Los criterios se aplican a los valores sin redondear.

Cada criterio se evalúa primero en cada participante. Si falta el valor de iCanClean, la comparación cuenta como no cumplida; si falta el del comparador, por ejemplo, porque el método quedó sin

configuración elegible, es no evaluable, porque esa falta no informa sobre iCanClean. Un criterio se cumple frente a un comparador, y el requisito previo frente a la identidad, si se cumple en al menos cuatro de los cinco participantes; si no alcanza ese número, pero lo alcanzaría contando las comparaciones no evaluables, queda indeterminado, y en otro caso no se cumple. La hipótesis recibe una de cuatro etiquetas:

- **respaldada**, si se cumplen el requisito previo y los criterios 1 y 2 frente a los tres comparadores;
- **parcialmente respaldada**, si se cumple el requisito previo y, frente a al menos dos comparadores, se cumplen los criterios 1 y 2;
- **no respaldada**, en cualquier otro caso en que la etiqueta no dependa de un criterio indeterminado, incluido aquel en que no se cumple el requisito previo;
- **indeterminada**, si la etiqueta cambia según que los criterios indeterminados se consideren cumplidos o no, caso en que se informa, además, el rango de etiquetas posibles.

Como la calibración de ASR es frágil (Sección 4.4.1), cada comparación entre iCanClean y ASR se recalcula con cada repetición de ASR, sin repetir la selección; es estable si su resultado, es decir, cumplimiento, empate o no cumplimiento, no cambia en las repeticiones cuya calibración no falló, y frágil si cambia en alguna. La etiqueta recalculada con cada repetición se informa junto con la principal, que no se reemplaza.

El requisito previo es poco exigente para iCanClean: como el método resta en cada ventana una proyección ortogonal, la energía de cada canal centrado en esa ventana no aumenta (Sección 3.5), y el requisito se cumple casi siempre que selecciona algún par, aunque las referencias no contengan información sobre el artefacto. Los criterios tampoco incluyen el control de especificidad, de modo que la etiqueta no distingue una atenuación que depende de las referencias propias de una que también se obtiene con las de otro participante. Por eso, la etiqueta se lee junto con los controles de especificidad y de actividad ocular, en particular con la fracción $\Delta\mathrm{RMS}_{\mathrm{mast}}(\mathrm{a})/\Delta\mathrm{RMS}_{\mathrm{mast}}(\mathrm{propias})$ de iCanClean, y resume la comparación según estos criterios: no califica la calidad de la limpieza.

La etiqueta tiene, además, sesgos en ambos sentidos, que se desprenden de lo observado en los datos de desarrollo. A favor de iCanClean, con umbrales bajos o ventanas breves, el método puede cumplir el requisito previo y el criterio de atenuación frente al AF por bloques y a la CCA con rezagos retirando actividad que también retira con las referencias de otro participante, algo que la etiqueta no detecta. En contra de iCanClean, como ASR puede retirar gran parte de la señal en masticación, mientras que iCanClean retira por ventana a lo sumo un par canónico por cada referencia, el criterio de atenuación frente a ASR es difícil de cumplir, de modo que, en la práctica, la etiqueta más alta alcanzable es «parcialmente respaldada».

La variante extendida no interviene en la etiqueta. Se considera que aporta si su S_{bal} medio, es decir, la media en los cinco participantes del S_{bal} de $\zeta^{*}(p)$ evaluado en p, supera en al menos 1 punto al de iCanClean, calculado del mismo modo; ambos valores se comparan sin redondear. Si falta el

valor de alguno de los dos métodos en algún participante, el criterio no se evalúa y se informa el motivo. Sus resultados se resumen en el Apéndice F.

4.5.9. Agregación y estadística descriptiva

La unidad de análisis es el participante. Para cada participante, método y condición, las métricas se promedian con igual peso sobre los diez canales primarios, y el resumen global es la media de los participantes, también con igual peso. Así, un participante con más ventanas no pesa más en el resultado, y las ventanas de un mismo participante no se tratan como observaciones independientes. Se informan los valores de cada participante, su media, su rango y un intervalo bootstrap percentil del 95 % para la media, con 2000 réplicas en que se remuestrean cinco participantes con reemplazo, con semilla 0. Como solo existen 126 remuestreos distintos, el intervalo es solo descriptivo e inestable, y no se interpreta como evidencia de generalización.

4.5.10. Análisis de sensibilidad

Los análisis de sensibilidad examinan cuánto dependen los resultados de decisiones de diseño. Se calculan con la configuración elegida de cada método, salvo el último, que repite la selección de iCanClean; no reemplazan la selección ni la etiqueta principales y se presentan en el Apéndice D:

- **Referencia EMG:** las cuatro referencias alternativas de la Sección 4.2, con los métodos calibrados recalibrados.
- **ASR:** actualización cada 32 muestras, zona de calibración desplazada 1, 2, 4 y 8 muestras hacia su guarda final y O1 retirado en lugar de O2; una repetición cuya calibración falla se registra como fallo.
- **Ventanas independientes del diseño:** el criterio 1 y el requisito previo se calculan solo con las 47 ventanas de masticación cuyos tramos no intervinieron en la definición del diseño (Sección 4.5.8), de E1M2_R2 y de E2.
- **Cercanía a la reserva:** sin la primera ventana evaluada de cada pieza que contiene una reserva, a solo 4 s de la calibración.
- **Última ventana de iCanClean con 4 s:** sin la última ventana evaluada de las piezas en que esa ventana se extiende hacia la guarda final.
- **Canales de gran amplitud:** con el promedio de referencia calculado sin los canales marcados, también en las calibraciones.
- **Duración de la ventana de iCanClean:** la selección de iCanClean se repite con ventanas de 2, 4 y 8 s, y las de 8 s se alinean como las de 4 s. Con 8 s, las tres ventanas de masticación de E1M1_R2 no son evaluables, porque sus piezas, de 6,9 y 9,1 s, son más breves que la guarda

inicial y una ventana, y la media se calcula con los otros cuatro participantes. La lectura es descriptiva: una duración con mejores métricas no reemplaza la selección principal.

4.6. Visualización: inspección visual

La inspección visual en los dominios temporal, espectral y tiempo-frecuencia complementa las métricas, porque permite identificar modificaciones de la señal que una única métrica podría no reflejar. Se hace en reposo y en masticación, para distinguir una modificación general de la señal de una reducción localizada del artefacto: una salida de menor amplitud puede parecer favorable en el dominio temporal, pero debe contrastarse con R^2_{reposo}, CC, la densidad espectral de potencia (PSD) y el espectrograma.

Para la inspección se seleccionan los canales AF3 y AF4, que forman parte de los diez canales primarios. Están uno a cada lado de la región frontal anterior, próximos a la distribución frontal asociada a la actividad facial y mandibular, y permiten comparar ambos hemisferios.

Las figuras de ejemplo se eligen con una regla que no depende de los resultados: el participante es E1M1_R2, con la configuración $\zeta^*(\text{E1M1_R2})$ de cada método; las trazas temporales muestran la primera ventana de evaluación de cada condición, y la PSD se calcula con todas. El espectrograma cubre las primeras cinco ventanas, es decir, 10 s, de la zona de evaluación más larga de cada condición, sin cruzar un corte; en la masticación de E1M1_R2, esa zona contiene dos ventanas, y su espectrograma abarca 4 s. Como corresponde a un solo participante y a dos canales, la inspección ilustra el comportamiento de cada método en E1M1_R2 y no se generaliza a los demás participantes.

El espectrograma se estima mediante la transformada de Fourier de tiempo corto de SciPy, con ventana de Hamming de 2048 muestras (2 s), solapamiento de 1920 muestras y transformada de 2048 puntos, sin extender la señal en los bordes, con una resolución de 0,5 Hz y un paso temporal de 125 ms. La densidad se expresa en decibeles respecto de $1\,\mu\text{V}^2/\text{Hz}$, en 1–45 Hz, con una escala de color común a la entrada y a la salida, fijada entre los percentiles 20 y 99,5 de ambos espectrogramas.

4.7. Prueba semisintética

Las métricas de la Sección 4.5 comparan la salida con el EEG observado. Para disponer de un objetivo conocido se construye, además, una prueba semisintética: como en los conjuntos semisintéticos en que una señal de referencia se contamina de forma aditiva con un nivel definido mediante valores RMS [27], a un reposo observado se le suma la contaminación de masticación del mismo participante con una amplitud controlada. Para cada participante, $\boldsymbol{X}_{\text{rep}}$ es la parte de su pieza de reposo de evaluación que sigue a la reserva, y $\boldsymbol{X}_{\text{mast}}$, una pieza de masticación de evaluación, ambas con sus guardas y recortadas a la duración de la más corta; $\boldsymbol{X}_{\text{mast}}$ se centra restando su media. La mezcla es

$$\boldsymbol{X}_{\text{mezcla}} = \boldsymbol{X}_{\text{rep}} + \kappa\, \boldsymbol{X}_{\text{mast}}, \tag{4.10}$$

donde el factor κ se elige de modo que la relación señal-ruido

$$\mathrm{SNR} = 20\log_{10}\frac{\mathrm{RMS}(\boldsymbol{X}_{\mathrm{rep}})}{\mathrm{RMS}(\kappa\boldsymbol{X}_{\mathrm{mast}})} \tag{4.11}$$

tome los valores -5, 0 y 5 dB. Esta definición corresponde a una razón de potencias en decibeles; EEGdenoiseNet, en cambio, usa $10\log_{10}$ del cociente de RMS [27], de modo que los niveles de ambos trabajos no son directamente comparables. El factor κ es uno solo por participante y nivel de SNR, calculado con la RMS del conjunto de las zonas de evaluación de cada condición en los diez canales primarios, por lo que la SNR de cada mezcla es aproximada. Las referencias EMG de la pieza de masticación se multiplican por el mismo factor, y se forma una mezcla por cada pieza de masticación de evaluación y cada nivel de SNR.

Cada mezcla se procesa como una pieza, con las configuraciones $\zeta^{*}(p)$, sin reajustarlas, y con las calibraciones de la Sección 4.3. La salida $y[n]$ se compara en las ventanas de evaluación con el canal correspondiente $x_{\mathrm{rep}}[n]$ de $\boldsymbol{X}_{\mathrm{rep}}$, con las métricas que utiliza EEGdenoiseNet [27], promediadas sobre los diez canales primarios: el error relativo temporal y el espectral, este con las densidades de Welch en 1–45 Hz,

$$\mathrm{RRMSE}_{\mathrm{t}} = \frac{\|y - x_{\mathrm{rep}}\|}{\|x_{\mathrm{rep}}\|}, \qquad \mathrm{RRMSE}_{\mathrm{f}} = \frac{\|\widehat{P}_{y} - \widehat{P}_{\mathrm{rep}}\|}{\|\widehat{P}_{\mathrm{rep}}\|}, \tag{4.12}$$

y la correlación CC entre $y[n]$ y $x_{\mathrm{rep}}[n]$. La identidad se incluye como referencia, porque su diferencia con $\boldsymbol{X}_{\mathrm{rep}}$ corresponde a la contaminación agregada.

La contaminación $\boldsymbol{X}_{\mathrm{mast}}$ contiene también la actividad cerebral del tramo de masticación, y $\boldsymbol{X}_{\mathrm{rep}}$ es un reposo observado que no está libre de artefactos. Por lo tanto, $\boldsymbol{X}_{\mathrm{rep}}$ es un objetivo conocido para la mezcla, pero no una señal cerebral pura: la prueba evalúa la capacidad de recuperar $\boldsymbol{X}_{\mathrm{rep}}$, y no la separación de la actividad cerebral. Además, las referencias de la mezcla describen solo la contaminación agregada y no la actividad muscular residual del reposo, lo que favorece a los métodos con referencia frente a ASR, que puede corregir también esa actividad.

CAPÍTULO

5

Resultados

En este capítulo se presentan los resultados obtenidos al aplicar el protocolo descrito en el Capítulo 4 a los métodos de limpieza comparados, ASR, AF por bloques, CCA con rezagos e iCanClean, junto con la identidad, que sirve como línea base. La variante extendida propia (Sección 3.6) se evaluó con el mismo procedimiento, pero su S_{bal} medio, con la configuración elegida para cada participante, fue de 51,6, menor que el de iCanClean, de 55,1 (diferencia de $-3{,}47$ puntos), y no alcanzó la ventaja mínima de un punto del criterio de la Sección 4.5.8. Sus resultados se resumen en el Apéndice F. Después de los resultados de cada método, se comparan los métodos entre sí y con la identidad, se presenta la evaluación semisintética y se discuten los resultados.

5.1. Ventanas de evaluación

La evaluación comprende 52 ventanas de reposo y 126 ventanas de masticación, distribuidas estas últimas en 8 piezas (Tabla 5.1). El número de ventanas difiere sobre todo en masticación: E1M2_R2 aporta 54, mientras que E1M1_R2 conserva solo 3, repartidas en 2 piezas, porque su intervalo de masticación contiene un corte y su primera parte forma la reserva de calibración (Sección 4.3).

La calibración de ASR conservó entre 10,4 s, en E2E3S7, y 19,4 s, en E2E3S9, después del descarte de las ventanas con amplitudes atípicas, que no depende de la tolerancia k (Tabla 5.1). Las verificaciones del preprocesamiento (Apéndice C.1) solo señalaron la continua residual de la pieza de masticación de E1M1_R2 comprendida entre la reserva y el corte, y canales de gran amplitud en las piezas de masticación de E1M2_R2.

	Reposo	Masticación		ASR
Participante	Ventanas	Piezas	Ventanas	Calibración efectiva (s)
E1M1_R2	8	2	3	17,4
E1M2_R2	11	3	54	13,5
E2E3S7	11	1	23	10,4
E2E3S8	11	1	23	12,8
E2E3S9	11	1	23	19,4
Total	52	8	126	—

Tabla 5.1: Ventanas de evaluación de cada condición, piezas de masticación evaluadas y duración de la calibración de ASR que se conserva después de descartar las ventanas con amplitudes atípicas (Sección 4.4.1), para cada participante. La duración de la calibración de ASR no se suma entre participantes (—).

5.2. Resultados de ASR

5.2.1. Selección de la configuración

La configuración elegida fue ASR-09 en 2 selecciones, las de E1M1_R2 y E2E3S8, y ASR-11 en las 3 restantes (Tabla B.1). Ambas tienen la menor fracción dimensional f_{ASR} y tolerancias contiguas: ASR-11, la mayor k de la grilla, y ASR-09, la inmediatamente inferior. En las selecciones para E1M1_R2, E1M2_R2 y E2E3S8, la elección la fijó el desempate (Tabla B.2). Con la tolerancia de LIT-ASR e HIS-ASR y con las mayores, ambos valores de f_{ASR} dieron las mismas métricas en todos los participantes; por ello, LIT-ASR e HIS-ASR, que difieren solo en f_{ASR}, coinciden entre sí y con configuraciones de la grilla.

En la Figura 5.1, en cada selección, una mayor $\Delta\mathrm{RMS}_{\mathrm{mast}}$ se acompaña de un menor R^2_{reposo}, y la configuración elegida se ubica cerca del extremo de mayor similitud en reposo, hacia el que crece k. Su valor en el participante evaluado, en cambio, se aleja de la media de los otros cuatro en direcciones distintas.

5.2.2. Desempeño cuantitativo

En masticación, ASR redujo la RMS en todos los participantes, aunque en magnitudes muy distintas: $\Delta\mathrm{RMS}_{\mathrm{mast}}$ fue de 85,8 % en E1M2_R2 y de solo 9,58 % en E2E3S7, con una media de 60,6 % (Tabla 5.2). La reducción de potencia en 20–45 Hz siguió el mismo orden, con una media de 65,9 %.

En reposo, la modificación también fue desigual: ASR dejó la señal de E2E3S7 prácticamente sin cambios y modificó poco la de E1M1_R2, mientras que en los demás participantes la modificó de forma apreciable, hasta un R^2_{reposo} de 23,2 % y un $Q_{\mathrm{RMS,reposo}}$ de 0,452 en E1M2_R2. La correlación con signo en reposo fue positiva en todos los participantes, de modo que los valores bajos de R^2_{reposo} no se deben a una inversión de la señal.

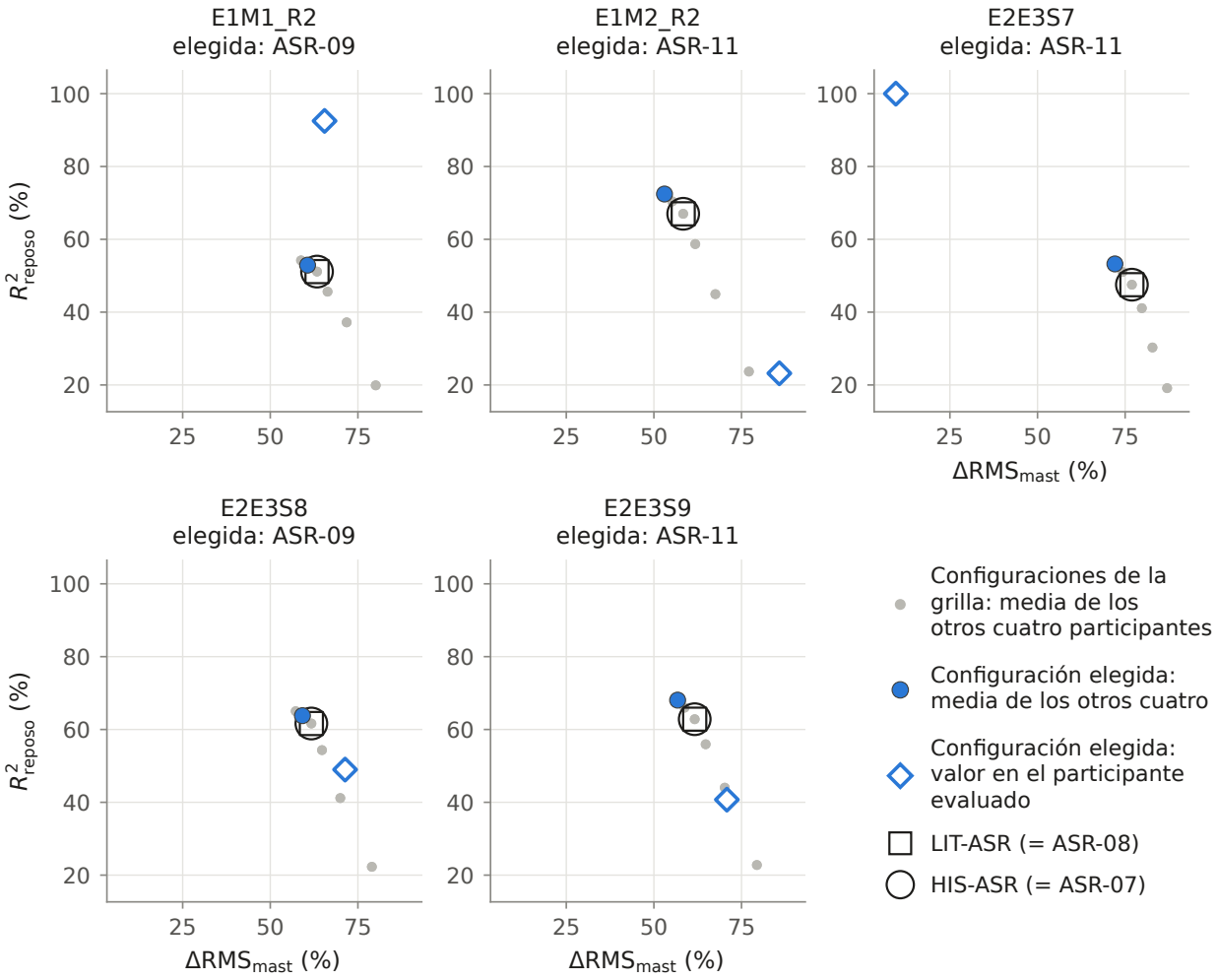


Figura 5.1: ASR: relación entre R^2_{reposo} y $\Delta\text{RMS}_{\text{mast}}$ de todas las configuraciones de la grilla. Cada panel corresponde a un participante evaluado, cuyo nombre y configuración elegida se indican sobre el panel. Los valores de la grilla se promedian en los otros cuatro participantes, con los que se elige su configuración, y se marcan la configuración elegida y su valor en el participante evaluado. Todos los paneles comparten escalas. Las configuraciones con métricas idénticas se superponen en un mismo punto, y la configuración elegida puede no ser la de mayor puntaje medio, porque entre las configuraciones empatadas decide $|1 - Q_{\text{RMS,reposo}}|$ (Sección 4.5.7). El cuadrado y el círculo abiertos ubican LIT-ASR e HIS-ASR, configuraciones informadas sin selección que coinciden con configuraciones de la grilla (Tabla B.4). La figura describe la exploración de la grilla y no informa desempeño.

	Reposo				Masticación				Puntaje	
Participante	R^2	CC	Q_{RMS}	E_{rel}	$\Delta\mathrm{RMS}$	Reducción en 20–45 Hz	CC	E_{rel}	S_{bal}	$S_{\mathrm{bal,banda}}$
E1M1_R2	92,5 %	0,962	0,964	0,259	65,5 %	65,8 %	0,210	0,979	79,0	79,1
E1M2_R2	23,2 %	0,420	0,452	0,862	85,8 %	86,8 %	0,0742	1,00	54,5	55,0
E2E3S7	100,0 %	1,00	1,00	0,0000	9,58 %	10,0 %	0,871	0,412	54,8	55,0
E2E3S8	49,0 %	0,690	0,730	0,699	71,3 %	85,2 %	0,194	0,983	60,2	67,1
E2E3S9	40,7 %	0,577	0,587	0,723	70,8 %	82,0 %	0,280	0,953	55,8	61,4
Media	61,1 %	0,730	0,747	0,509	60,6 %	65,9 %	0,326	0,865	60,8	63,5
Intervalo bootstrap 95 %, límite inferior	35,4 %	0,537	0,561	0,196	34,1 %	36,5 %	0,146	0,638	54,9	56,3
Intervalo bootstrap 95 %, límite superior	86,8 %	0,923	0,932	0,774	77,7 %	84,9 %	0,606	0,988	70,3	72,0

Tabla 5.2: ASR: métricas de cada participante, con la configuración elegida con los otros cuatro participantes, y su media. Los límites del intervalo bootstrap percentil del 95 % de la media son solo descriptivos (Sección 4.5.9). R^2, $\Delta\mathrm{RMS}$ y la reducción de potencia en 20–45 Hz se expresan en porcentaje. S_{bal} y $S_{\mathrm{bal,banda}}$ corresponden a las ecuaciones (4.7) y (4.8), y el S_{bal} de cada participante es el de la Tabla B.1.

En masticación, salvo en E2E3S7, la salida conservó poca relación lineal con la entrada, y la señal eliminada tuvo una amplitud cercana a la de la entrada. Esto es coherente con la reconstrucción de ASR, que reemplaza las direcciones que superan el umbral por una combinación de las conservadas (Sección 3.2), y no con la resta de una componente pequeña. Por lo tanto, lo eliminado incluye toda la actividad que la salida dejó de reproducir, que durante la masticación puede comprender también actividad cerebral. Lo mismo ocurrió, en menor medida, en el reposo de E2E3S8, E2E3S9 y E1M2_R2.

En la Figura 5.2, la pérdida de similitud en reposo de E1M2_R2, E2E3S8 y E2E3S9 se concentra en los canales frontopolares y frontales anteriores. En masticación, la reducción fue similar en todos los canales primarios de E1M2_R2, y en E2E3S7 se concentró en T7 y T8.

La media de ASR resume valores muy dispersos (Tabla 5.2). E1M1_R2, que combinó una reducción apreciable en masticación con un reposo poco modificado, obtuvo el mayor S_{bal}, de modo que la media, 60,8, supera el valor de los demás participantes, y su intervalo bootstrap descriptivo, de 54,9 a 70,3, es asimétrico respecto de ella.

En cuanto a su estabilidad, en E1M1_R2 y E2E3S8 estos valores dependen de detalles de la calibración que no forman parte de la configuración elegida: las repeticiones con la zona de calibración desplazada o con O1 retirado en lugar de O2 los modificaron de forma apreciable, sobre todo en reposo, mientras que en los demás participantes los modificaron poco (Apéndice D.1).

5.2.3. Inspección visual en AF3 y AF4

En reposo, la señal corregida casi no se distingue de la entrada en AF3 y AF4 (Figura 5.3); su PSD es menor en alfa y, en AF3, en las frecuencias más bajas, y los espectrogramas conservan los patrones de la entrada (Figura 5.4). Esto concuerda con el R^2_{reposo} de E1M1_R2, 92,5 %, uno de los participantes en que ASR menos modificó el reposo (Tabla 5.2).

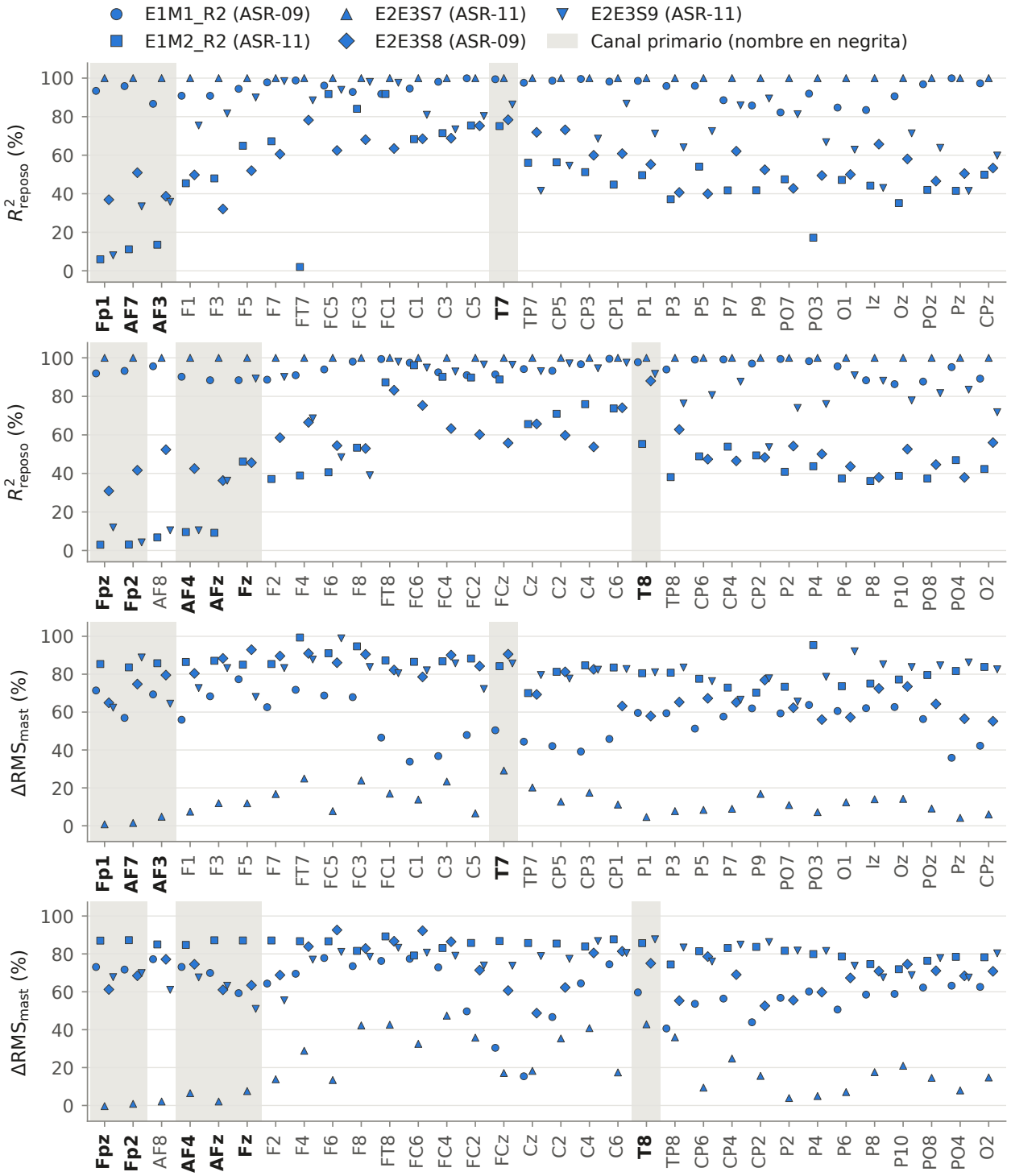


Figura 5.2: ASR: R^2_{reposo} (paneles superiores) y $\Delta\text{RMS}_{\text{mast}}$ (paneles inferiores) en cada canal EEG, con la configuración elegida para cada participante. Los canales siguen el orden de registro y, para cada métrica, se reparten en un par de paneles. Cada participante tiene un marcador propio, y la leyenda indica su configuración. Las franjas grises y los nombres en negrita señalan los diez canales primarios.

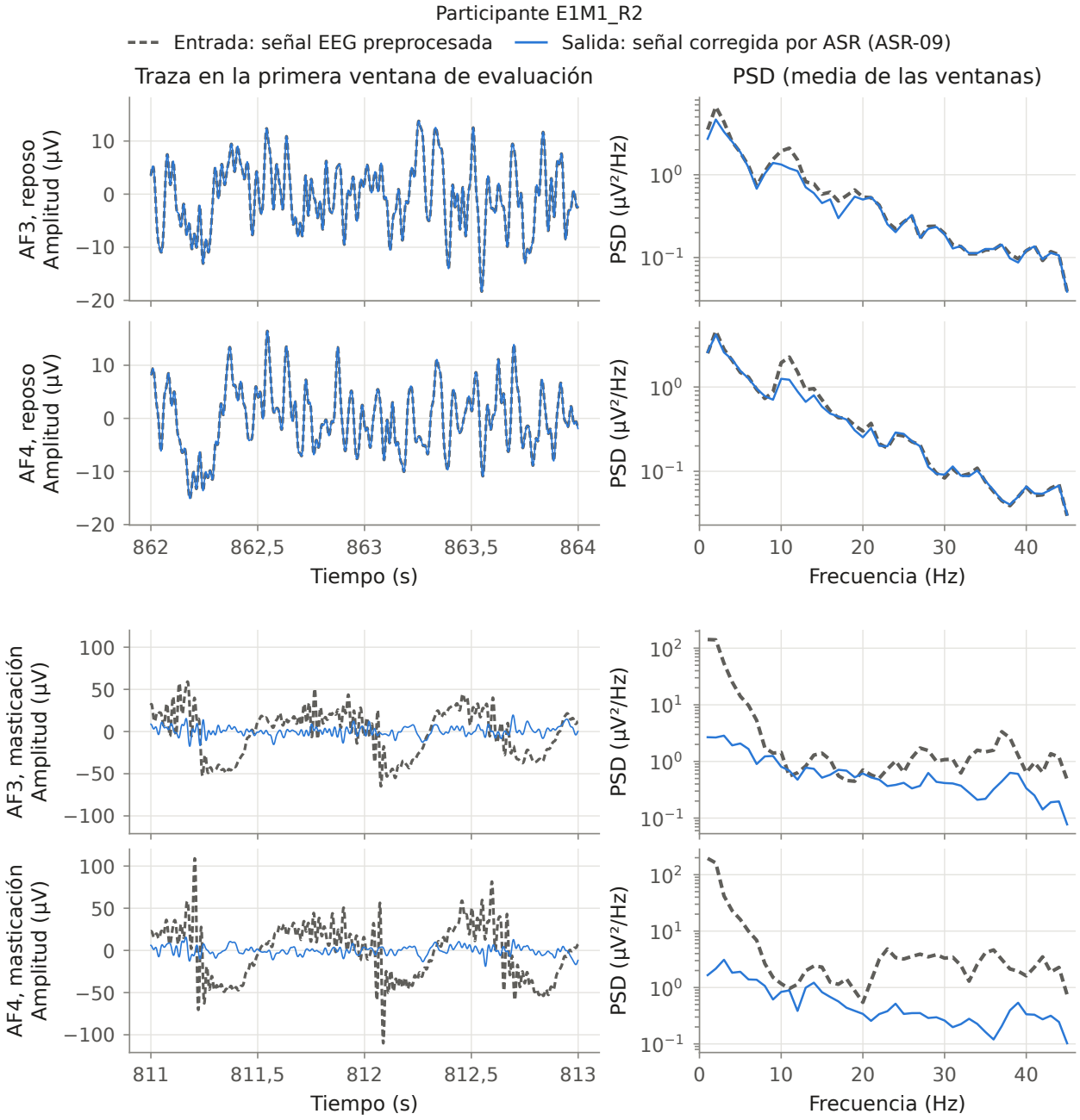


Figura 5.3: ASR: inspección visual en AF3 y AF4, participante E1M1_R2, con la configuración elegida para ese participante (Tabla B.1). Filas: AF3 y AF4 en reposo y en masticación. A la izquierda, la señal EEG preprocesada (entrada, trazo discontinuo) y la señal corregida (salida, trazo continuo) en la primera ventana de evaluación; a la derecha, su PSD promediada en todas las ventanas de evaluación de la condición. Donde la salida coincide con la entrada, el trazo discontinuo queda visible a ambos lados del continuo.

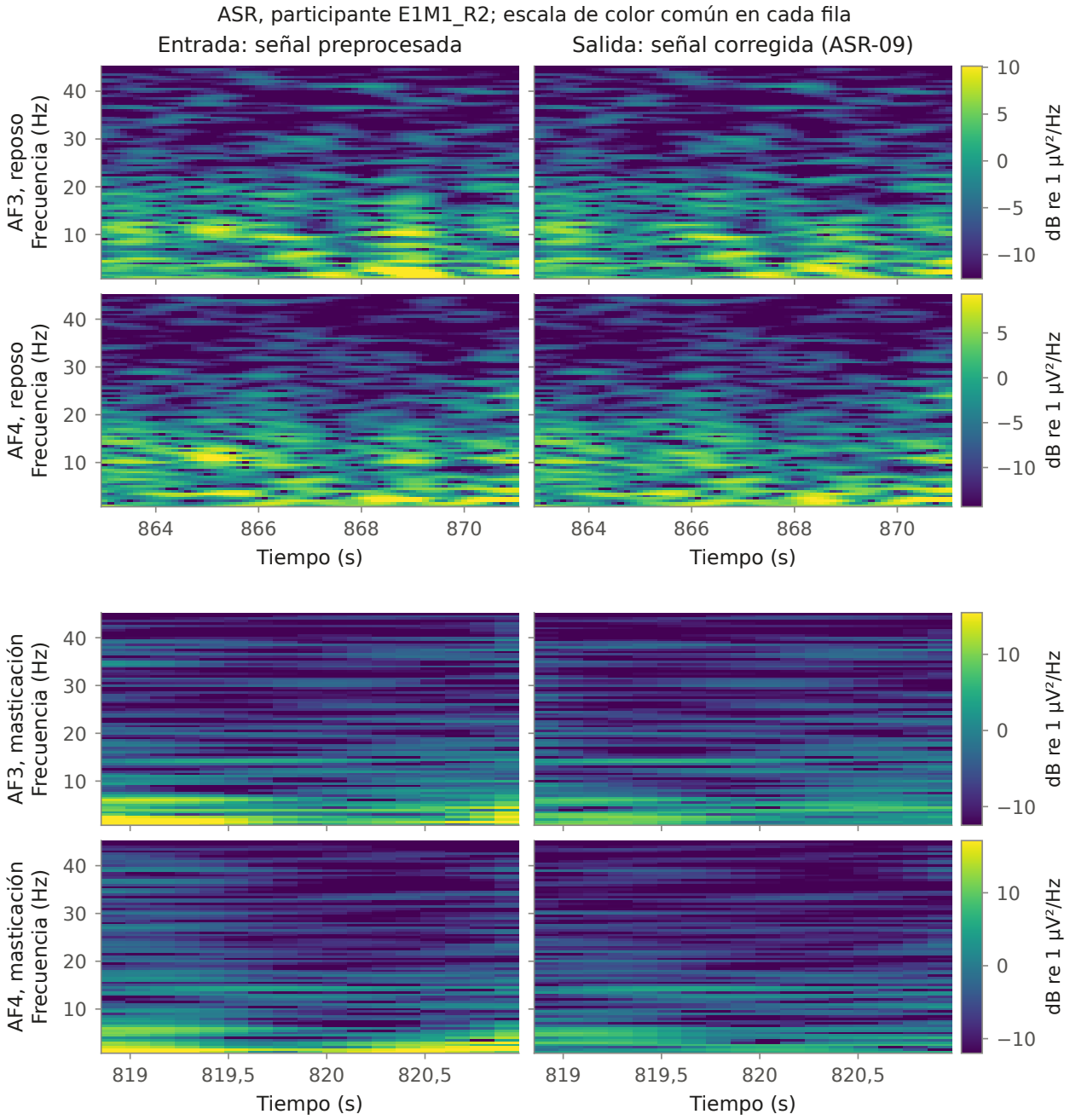


Figura 5.4: ASR: espectrogramas de la entrada (izquierda) y de la salida (derecha) en AF3 y AF4, participante E1M1_R2, en reposo y en masticación, con la misma configuración que la Figura 5.3. Cada espectrograma cubre el tramo que fija la regla de la Sección 4.6, es decir, las primeras ventanas de evaluación de la zona más larga de cada condición, o todas ellas si son menos, y la escala de color, en decibeles con la referencia de esa sección, es común a la entrada y la salida de cada fila.

En masticación, la señal corregida casi no reproduce las deflexiones lentas de gran amplitud de la entrada, y su potencia es menor en casi toda la banda, sobre todo en delta y θ y, de nuevo, en 20–45 Hz, donde la reducción de E1M1_R2 fue de 65,8 %. La inspección no permite atribuir esas deflexiones al artefacto muscular; su relación con la actividad ocular se examina con el control de la Sección 5.6.3.

5.3. Resultados del AF por bloques

5.3.1. Selección de la configuración

La configuración elegida fue AF-10 en las 5 selecciones (Tabla B.1). Incluye la envolvente en los regresores y combina la memoria mayor de la grilla, L_{AF}, con el menor paso de adaptación, μ. Solo en la selección para E2E3S7 la elegida no fue la de mayor puntaje medio, sino la que fijó el desempate (Tabla B.2).

En la Figura 5.5, las configuraciones se agrupan según el paso de adaptación: al aumentar μ, el R^2_{reposo} medio disminuye de forma marcada y la $\Delta\mathrm{RMS}_{\mathrm{mast}}$ media aumenta. La elegida tiene, en todos los paneles, el mayor R^2_{reposo} medio de la grilla, y HIS-AF se sitúa en el extremo de mayor $\Delta\mathrm{RMS}_{\mathrm{mast}}$ media, dentro del grupo de menor R^2_{reposo} medio. En el participante evaluado, el valor de la elegida se aparta de esa media, sobre todo en E2E3S7.

5.3.2. Desempeño cuantitativo

	Reposo				Masticación				Puntaje	
Participante	R^2	CC	Q_{RMS}	E_{rel}	$\Delta\mathrm{RMS}$	Reducción en 20–45 Hz	CC	E_{rel}	S_{bal}	$S_{\mathrm{bal,banda}}$
E1M1_R2	99,0 %	0,995	0,991	0,0980	19,0 %	1,77 %	0,931	0,379	59,0	50,4
E1M2_R2	98,0 %	0,990	0,977	0,143	18,6 %	1,01 %	0,899	0,443	58,3	49,5
E2E3S7	95,9 %	0,979	0,960	0,201	4,78 %	0,49 %	0,973	0,224	50,4	48,2
E2E3S8	99,4 %	0,997	0,991	0,0769	14,7 %	0,50 %	0,918	0,393	57,1	49,9
E2E3S9	96,2 %	0,981	0,985	0,194	13,3 %	1,89 %	0,929	0,349	54,8	49,0
Media	97,7 %	0,988	0,981	0,143	14,1 %	1,13 %	0,930	0,357	55,9	49,4
Intervalo bootstrap 95 %, límite inferior	96,4 %	0,982	0,969	0,0986	9,33 %	0,60 %	0,912	0,286	53,0	48,8
Intervalo bootstrap 95 %, límite superior	99,0 %	0,995	0,990	0,187	18,0 %	1,69 %	0,954	0,411	58,3	50,0

Tabla 5.3: AF por bloques: métricas de cada participante, con la configuración elegida con los otros cuatro participantes, su media y los límites del intervalo bootstrap, como en la Tabla 5.2.

En reposo, el AF por bloques modificó poco el EEG observado, con un R^2_{reposo} medio de 97,7 % (Tabla 5.3). Sin embargo, $Q_{\mathrm{RMS,reposo}}$ fue menor que uno en todos los participantes, con una media de 0,981. Es decir, el filtro atenuó levemente la amplitud del reposo, un cambio que R^2 no registra porque es invariante a la escala.

En masticación, $\Delta\mathrm{RMS}_{\mathrm{mast}}$ tuvo una media de 14,1 %, mientras que la reducción de potencia en 20–45 Hz fue pequeña en todos los participantes, con una media de 1,13 %. Por lo tanto, la

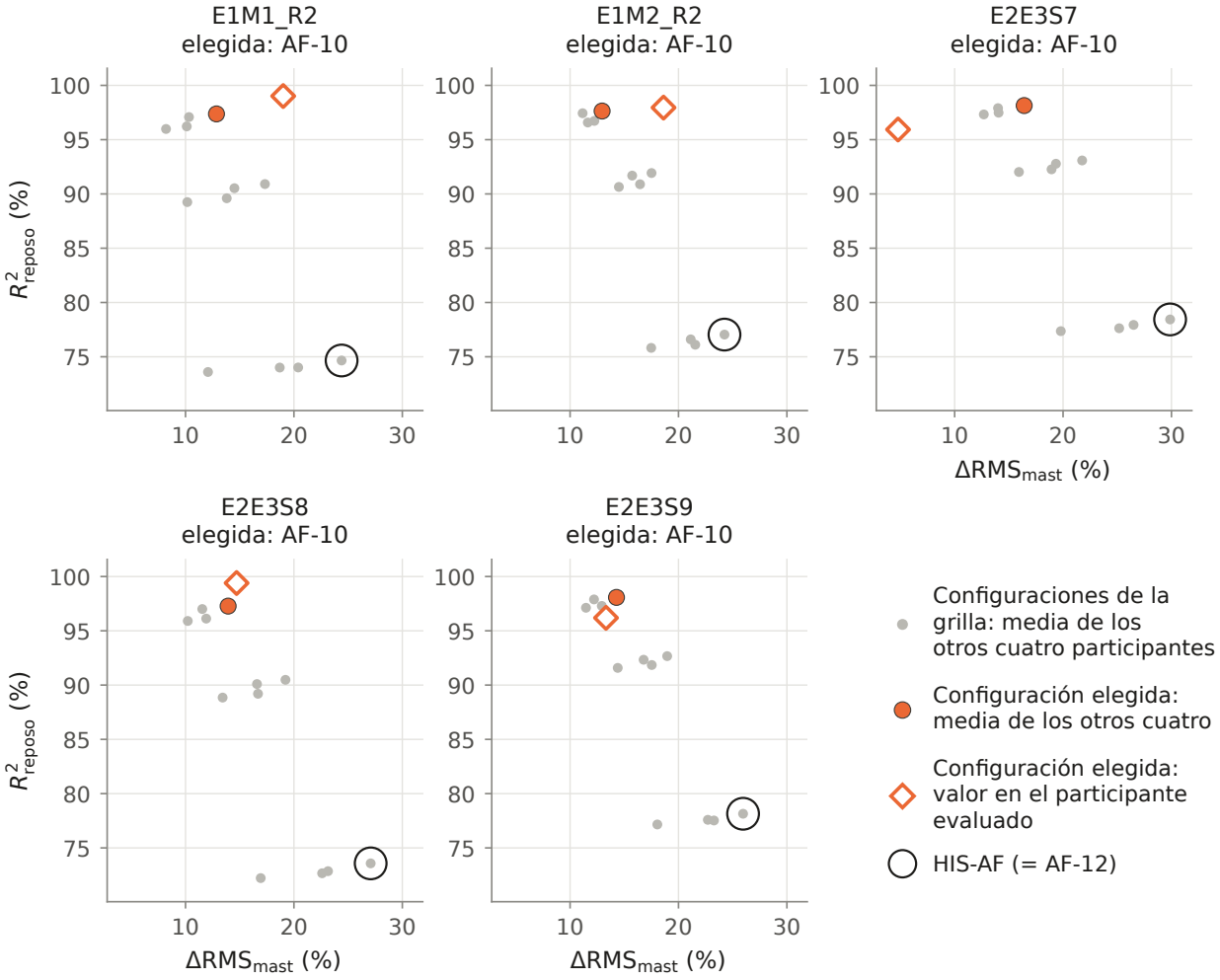


Figura 5.5: AF por bloques: relación entre R^2_{reposo} y $\Delta\text{RMS}_{\text{mast}}$ de todas las configuraciones de la grilla, como en la Figura 5.1. El círculo abierto ubica HIS-AF, que coincide con la configuración de la grilla con envolvente, la memoria mayor y el mayor paso de adaptación (Tabla B.4).

disminución de RMS se produjo principalmente fuera de esa banda (Sección 5.6.2). Además, E_{rel} superó en todos los participantes a la reducción relativa de RMS, de modo que la salida no es una copia escalada de la entrada. Salvo en E2E3S7, $Q_{\text{RMS,reposo}}$ fue claramente mayor que la razón de RMS en masticación, por lo que la reducción no corresponde a una atenuación uniforme de toda la señal.

La Figura 5.6 muestra que, en cada participante, R^2_{reposo} varía poco entre los canales primarios, mientras que $\Delta\text{RMS}_{\text{mast}}$ varía mucho, y el canal con la mayor reducción cambia de un participante a otro.

La configuración elegida no usa el mayor paso de adaptación, con el que la atenuación en 1–8 Hz es comparable a la del término de sesgo solo (Sección 4.5.7); el efecto de ese término se examina con el control (b) (Sección 5.6.3). Como la configuración elegida fue la misma en todas las selecciones,

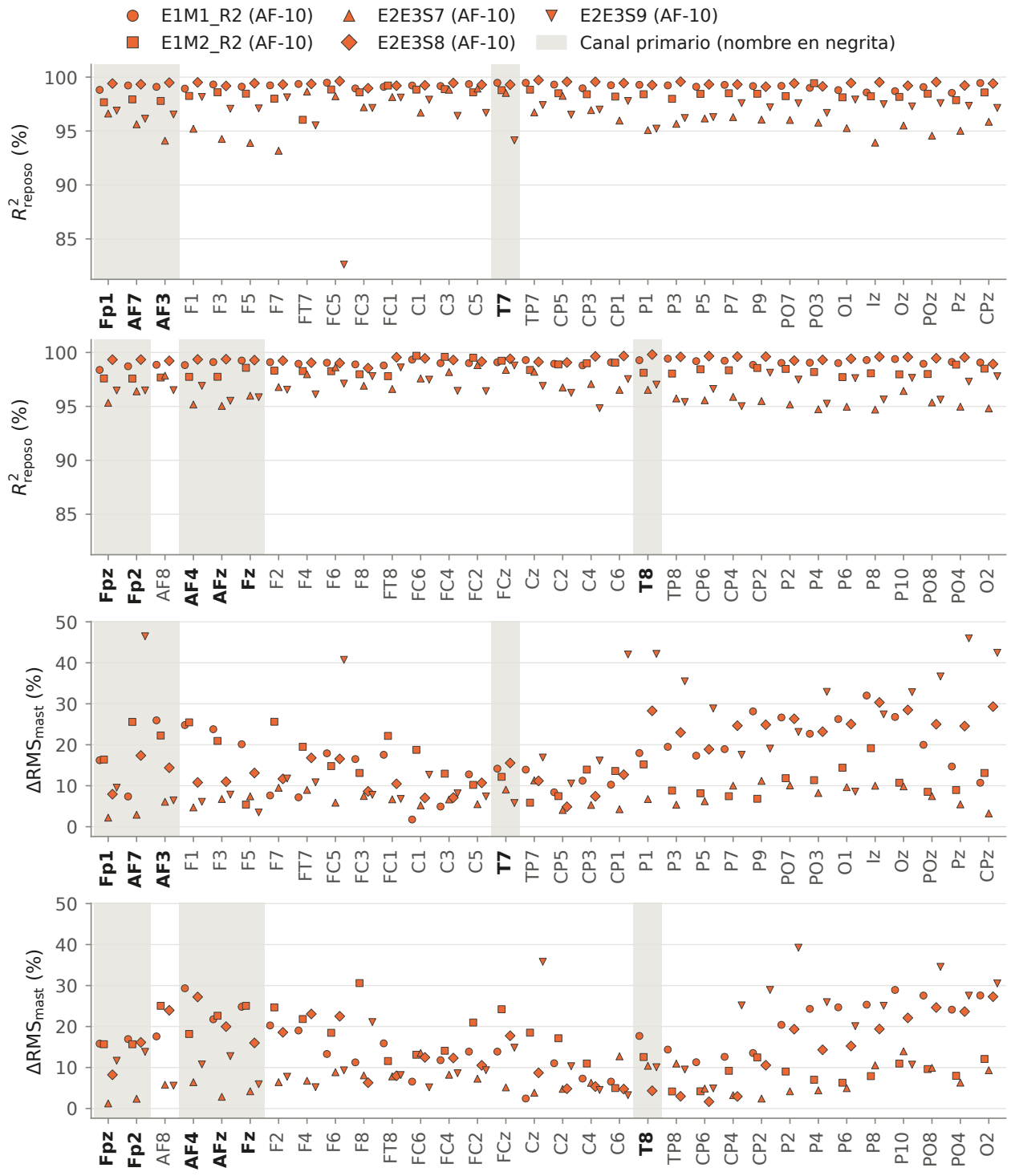


Figura 5.6: AF por bloques: R^2_{reposo} (paneles superiores) y $\Delta\text{RMS}_{\text{mast}}$ (paneles inferiores) en cada canal EEG, con la configuración elegida para cada participante, como en la Figura 5.2.

la dispersión entre participantes, con E2E3S7 en el extremo inferior de S_{bal}, R^2_{reposo}, $\Delta\text{RMS}_{\text{mast}}$ y $Q_{\text{RMS,reposo}}$, refleja diferencias entre ellos y no cambios de configuración.

5.3.3. Inspección visual en AF3 y AF4

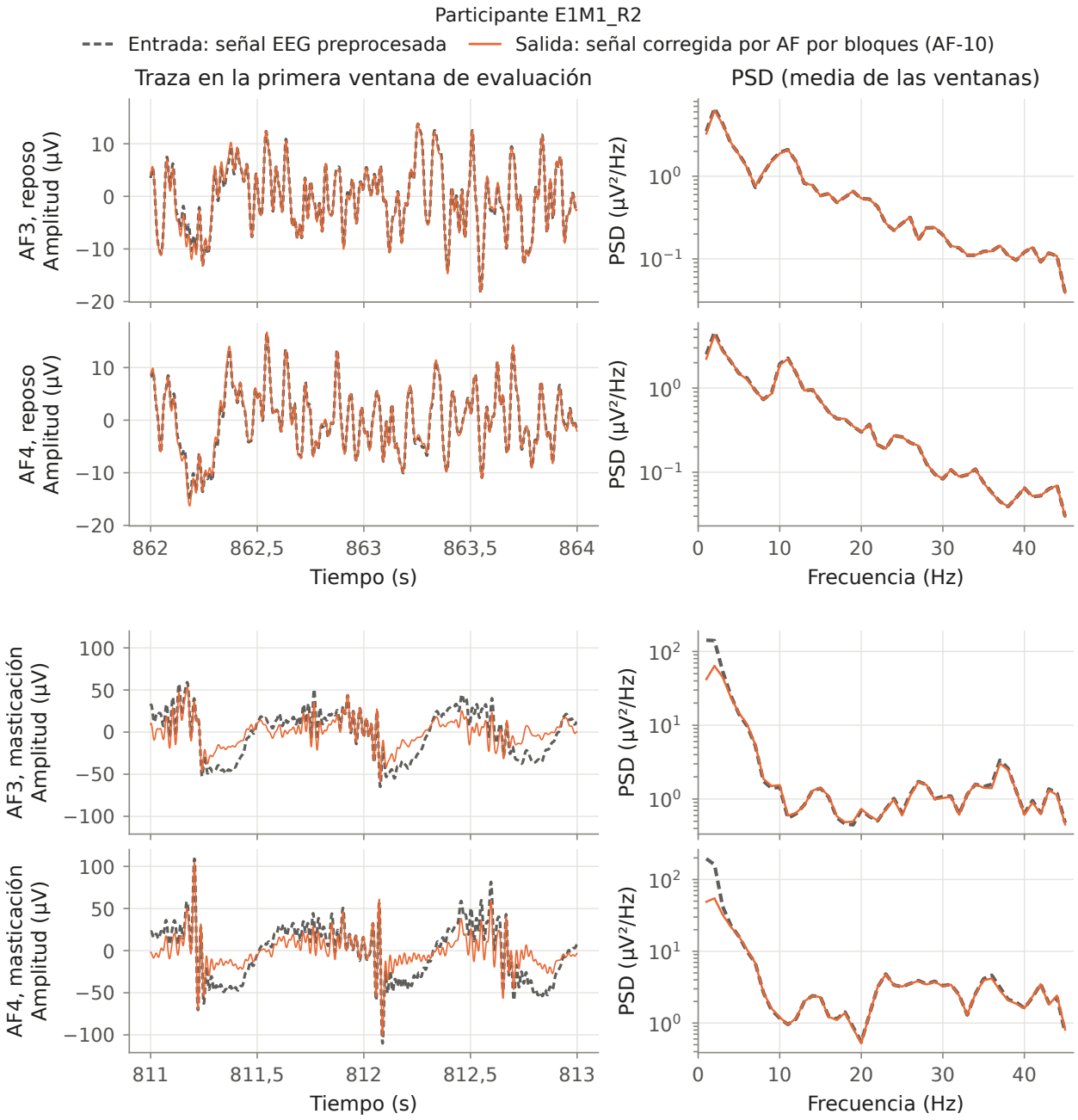


Figura 5.7: AF por bloques: inspección visual en AF3 y AF4, participante E1M1_R2, con la configuración elegida para ese participante (Tabla B.1). Filas, paneles y trazos como en la Figura 5.3.

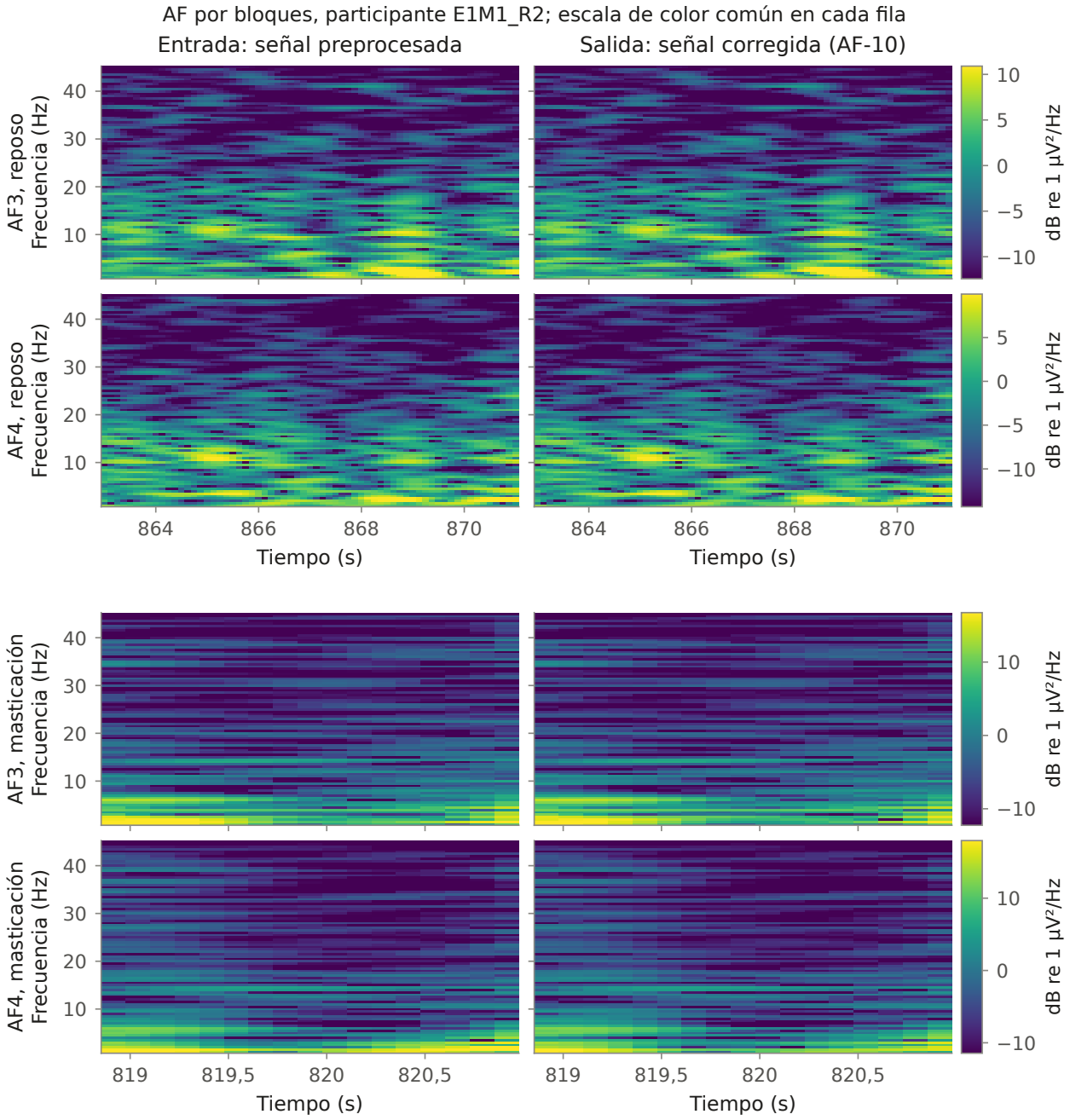


Figura 5.8: AF por bloques: espectrogramas de la entrada (izquierda) y de la salida (derecha) en AF3 y AF4, participante E1M1_R2, en reposo y en masticación, con la misma configuración que la Figura 5.7. Tramo y escala de color como en la Figura 5.4.

En reposo, la salida del AF por bloques casi coincide con la entrada en AF3 y AF4, en las trazas, las PSD y los espectrogramas (Figuras 5.7 y 5.8), de acuerdo con el R^2_{reposo} de E1M1_R2, 99,0 %. En masticación, la salida no reproduce las variaciones lentas de la entrada, y la diferencia de PSD se concentra en delta; por encima de esa banda, ambas curvas casi coinciden.

En la masticación de E1M1_R2, el cambio de potencia en delta fue de $-2{,}30$ dB con las referencias propias, frente a $-0{,}48$ dB con las de otro participante, $-0{,}36$ dB con la calibración rehecha con ellas y $-0{,}06$ dB con referencias nulas. Así, los controles reproducen solo una parte menor de esa atenuación (Sección 5.6.3).

5.4. Resultados de la CCA con rezagos

5.4.1. Selección de la configuración

La configuración elegida fue CCA-17 en 4 de las 5 selecciones y CCA-16 en la de E2E3S7 (Tabla B.1). Ambas usan el mayor rezago de la grilla y difieren solo en el umbral τ (Tabla B.3). En todas las selecciones, la elección dependió del desempate entre configuraciones empatadas (Tabla B.2).

En la Figura 5.9, los valores de la grilla, calculados como la media en los otros cuatro participantes, se agrupan en intervalos estrechos de ambas métricas, y la configuración elegida tiene en todos los paneles el mayor R^2_{reposo} medio. Su valor en el participante evaluado, en cambio, se aleja de ese grupo, de modo que, en estos datos, las diferencias entre participantes superan a las diferencias entre configuraciones.

5.4.2. Desempeño cuantitativo

	Reposo				Masticación				Puntaje	
Participante	R^2	CC	Q_{RMS}	E_{rel}	ΔRMS	Reducción en 20–45 Hz	CC	E_{rel}	S_{bal}	$S_{\text{bal,banda}}$
E1M1_R2	99,8 %	0,999	0,998	0,0359	19,6 %	6,82 %	0,838	0,531	59,7	53,3
E1M2_R2	97,9 %	0,989	0,990	0,140	2,90 %	−8,73 %	0,972	0,228	50,4	44,6
E2E3S7	96,8 %	0,984	0,972	0,172	2,94 %	4,87 %	0,972	0,215	49,9	50,8
E2E3S8	100,0 %	1,00	0,999	0,0096	17,4 %	3,29 %	0,831	0,533	58,7	51,6
E2E3S9	99,9 %	0,999	0,999	0,0316	10,1 %	14,7 %	0,901	0,394	55,0	57,3
Media	98,9 %	0,994	0,992	0,0779	10,6 %	4,18 %	0,903	0,380	54,7	51,5
Intervalo bootstrap 95 %, límite inferior	97,6 %	0,988	0,981	0,0237	4,37 %	−2,90 %	0,848	0,258	51,2	47,7
Intervalo bootstrap 95 %, límite superior	99,9 %	1,00	0,999	0,133	16,8 %	10,4 %	0,958	0,504	58,3	54,9

Tabla 5.4: CCA con rezagos: métricas de cada participante, con la configuración elegida con los otros cuatro participantes, su media y los límites del intervalo bootstrap, como en la Tabla 5.2.

En reposo, la CCA con rezagos modificó poco el EEG observado en E1M1_R2, E2E3S8 y E2E3S9, y más en E1M2_R2 y E2E3S7 (Tabla 5.4). En E2E3S8, el R^2_{reposo} redondeado no indica una salida idéntica a la entrada, como muestra su E_{rel}. $Q_{\text{RMS,reposo}}$ fue menor que uno en todos los participantes,

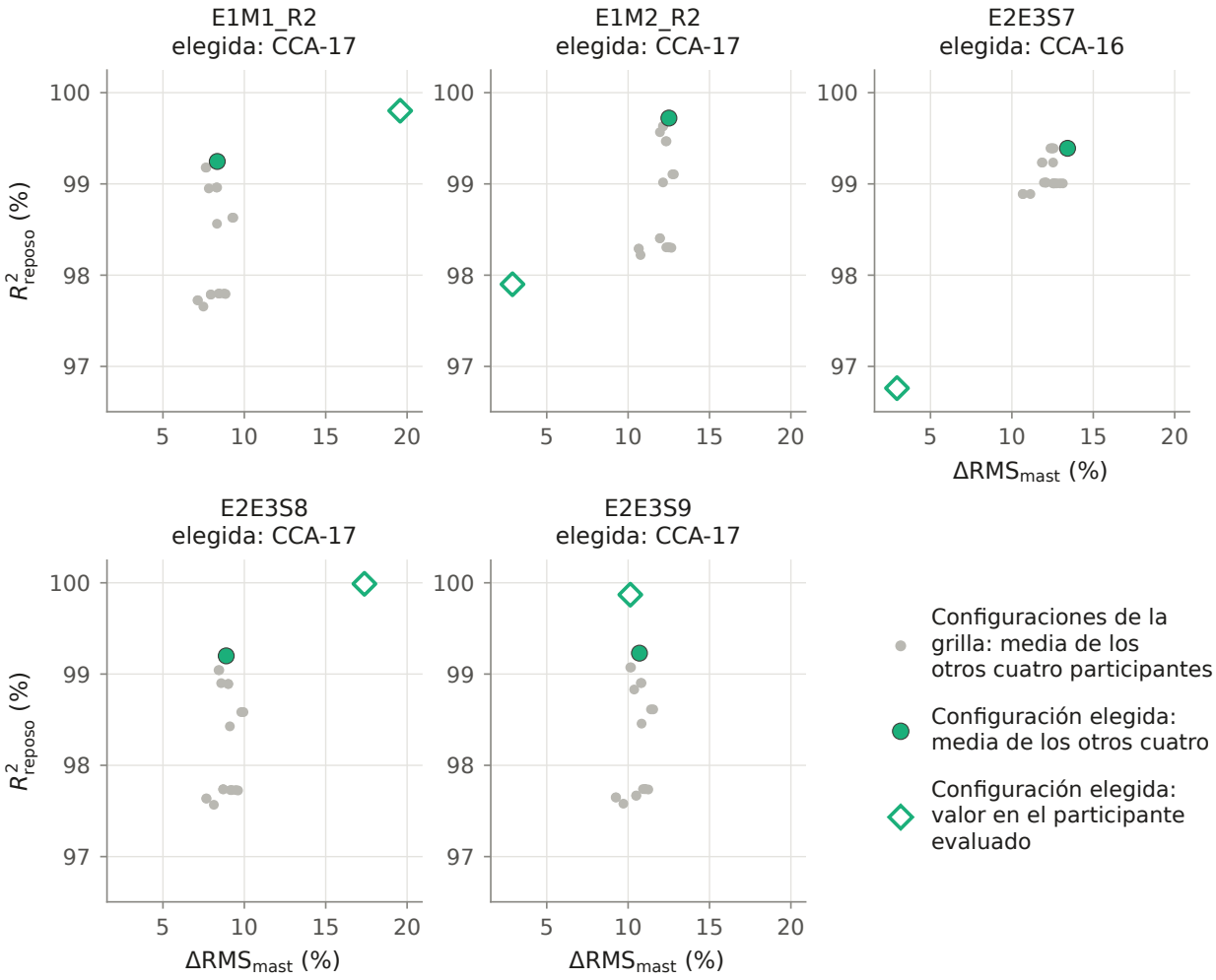


Figura 5.9: CCA con rezagos: relación entre R^2_{reposo} y $\Delta\text{RMS}_{\text{mast}}$ de todas las configuraciones de la grilla, como en la Figura 5.1. HIS-CCA queda fuera de la grilla y no se muestra (Tabla B.4).

con una media de 0,992. Además, en E1M2_R2 la potencia en 20–45 Hz aumentó en reposo, con una reducción de $-9{,}86\,\%$, un cambio que R^2_{reposo} resume solo en parte.

En masticación, $\Delta\text{RMS}_{\text{mast}}$ varió mucho entre participantes: fue mayor en E1M1_R2 y E2E3S8 y pequeña en E1M2_R2 y E2E3S7, con una media de 10,6 %. La reducción en 20–45 Hz tuvo una media de 4,18 % y fue negativa en E1M2_R2, donde la potencia de esa banda aumentó también en masticación. Entre participantes, la mayor reducción en masticación no se acompañó de una mayor modificación del reposo: E1M1_R2 y E2E3S8 combinan las mayores reducciones con valores altos de R^2_{reposo}, y E1M2_R2 y E2E3S7, las menores reducciones con los menores R^2_{reposo}.

En la Figura 5.10, R^2_{reposo} se mantiene cerca de su máximo en casi todos los canales de E1M1_R2, E2E3S8 y E2E3S9, y es menor en muchos canales de E1M2_R2 y E2E3S7. $\Delta\text{RMS}_{\text{mast}}$ es pequeña en todos los canales de E1M2_R2, y, fuera de los canales primarios, varias de las mayores reducciones se observan en canales parietales y occipitales.

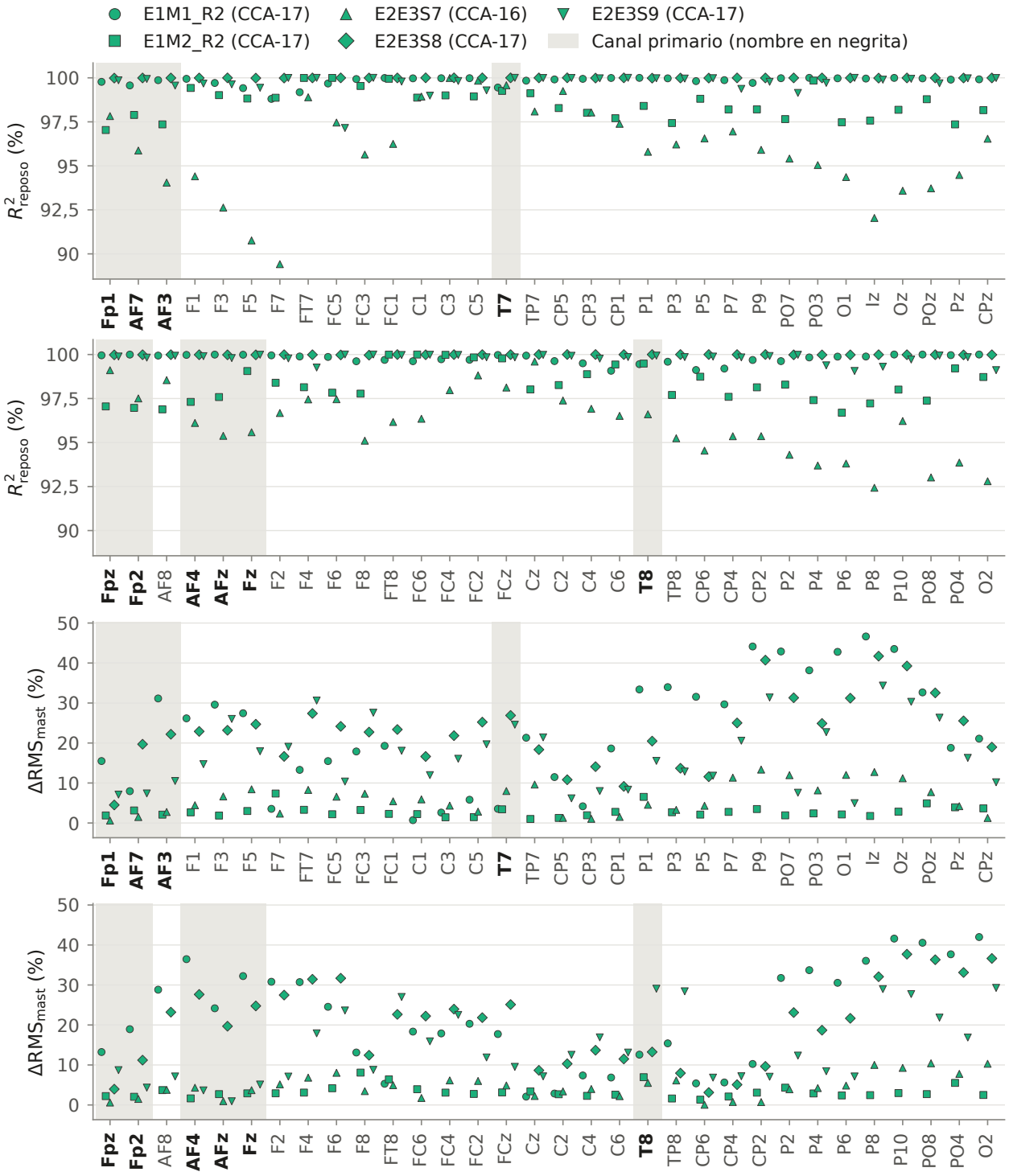


Figura 5.10: CCA con rezagos: R^2_{reposo} (paneles superiores) y $\Delta\text{RMS}_{\text{mast}}$ (paneles inferiores) en cada canal EEG, con la configuración elegida para cada participante, como en la Figura 5.2.

5.4.3. Inspección visual en AF3 y AF4

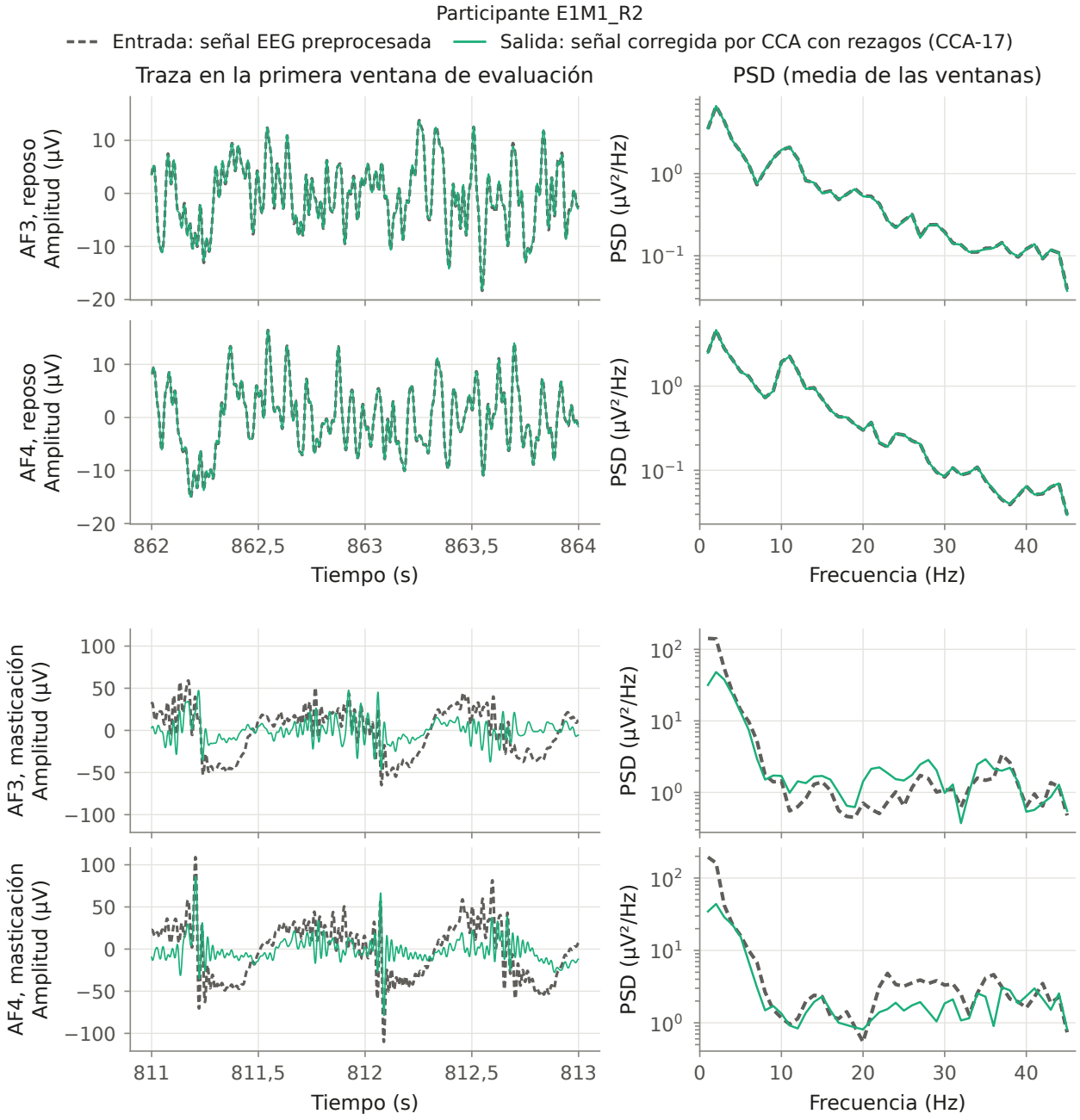


Figura 5.11: CCA con rezagos: inspección visual en AF3 y AF4, participante E1M1_R2, con la configuración elegida para ese participante (Tabla B.1). Filas, paneles y trazos como en la Figura 5.3.

En reposo, las trazas, las PSD y los espectrogramas de la entrada y de la salida de la CCA con rezagos prácticamente coinciden en AF3 y AF4 (Figuras 5.11 y 5.12), de forma coherente con el R^2_{reposo} de E1M1_R2, 99,8 %. En masticación, la salida no conserva en ningún canal las variaciones

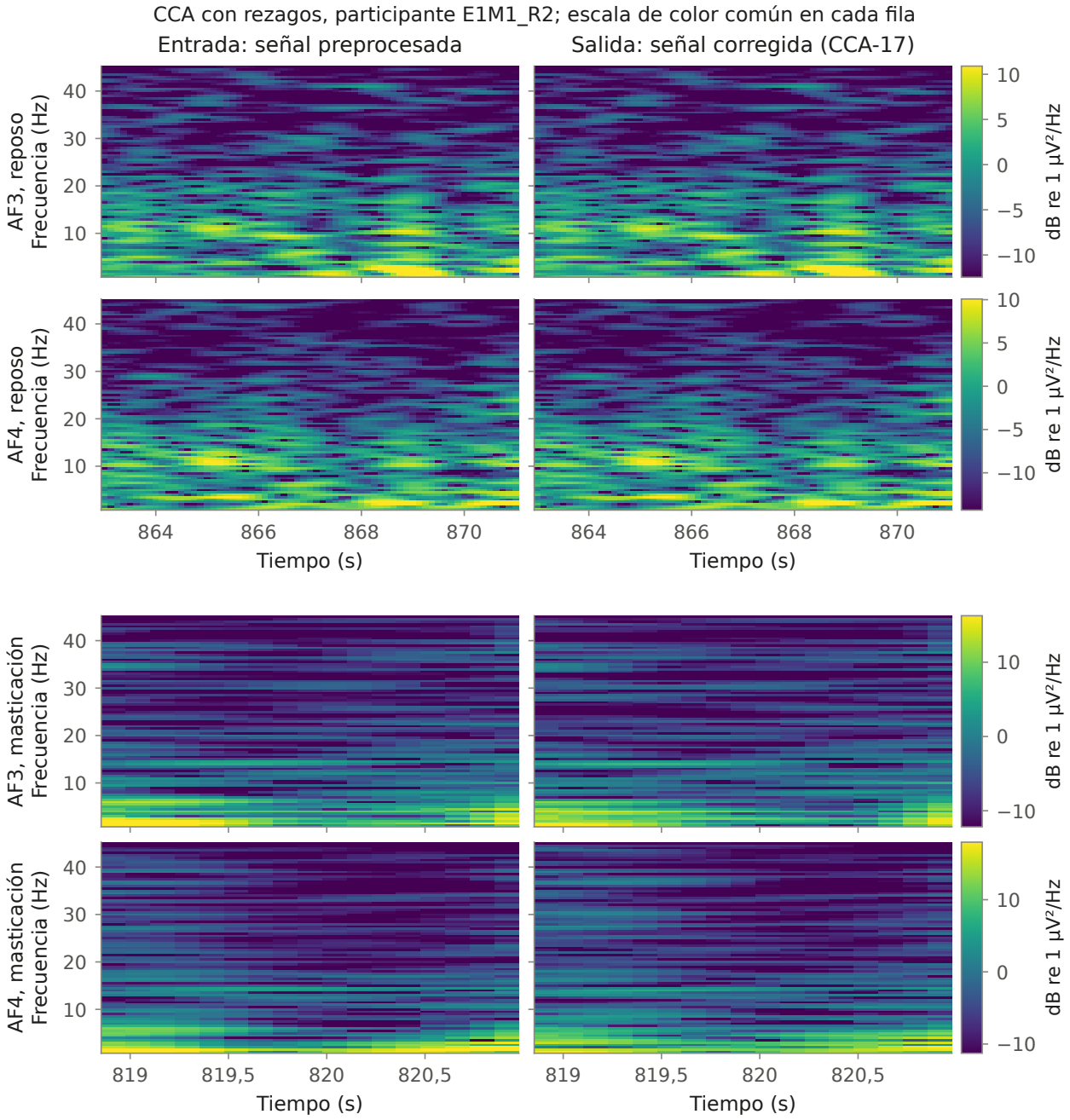


Figura 5.12: CCA con rezagos: espectrogramas de la entrada (izquierda) y de la salida (derecha) en AF3 y AF4, participante E1M1_R2, en reposo y en masticación, con la misma configuración que la Figura 5.11. Tramo y escala de color como en la Figura 5.4.

lentas de la entrada, y su PSD muestra menos potencia en las frecuencias más bajas. Por encima de θ, en cambio, el efecto difiere entre canales: la PSD de la salida supera a la de la entrada en buena parte de alfa y beta en AF3, y queda por debajo de ella en la mayor parte de beta y gamma en AF4. La inspección no permite establecer si la componente lenta retirada corresponde al artefacto.

5.5. Resultados de iCanClean

En esta sección se presentan los resultados de iCanClean, el algoritmo publicado, que se comparan con los de ASR, el AF por bloques y la CCA con rezagos en los criterios con que se examina la hipótesis (Sección 4.5.8).

5.5.1. Selección de la configuración

En las 5 selecciones se eligió la misma configuración, ICC-15, con ventanas de 4 s, el mayor umbral θ de la grilla y la fuente $\boldsymbol{U}$ (Tablas B.1 y B.2). Es la misma que LIT-ICC-4 (Tabla B.4). Su puntaje medio con los otros cuatro participantes varió entre 54,1 y 56,5, según el participante dejado fuera.

En la Figura 5.13, la configuración elegida se sitúa entre las configuraciones de mayor R^2_{reposo} medio y de menor $\Delta\text{RMS}_{\text{mast}}$ media de la grilla: todas las configuraciones con mayor reducción media en masticación tienen menor similitud media en reposo.

El umbral elegido se lee junto con el nivel que alcanza ρ^2_1 sin relación con el artefacto, medido con las referencias de otro participante en las ventanas del control (a) (Tabla D.7). Con ventanas de 4 s, ese umbral superó el percentil 95 de ese nivel en reposo, 0,606, y prácticamente coincidió con él en masticación, 0,651, mientras que los umbrales menores de la grilla quedaron bajo él; con ventanas de 2 s, incluso el mayor umbral de la grilla quedó bajo su mediana (Apéndice D.4). Con el umbral elegido, la selección con referencias de otro participante fue poco frecuente, con una fracción de 0,0625 de las ventanas en masticación y de 0,0357 en reposo, mientras que con las referencias propias fue frecuente en masticación (Tabla D.7).

Según la Tabla 5.5, con las referencias propias la fracción de ventanas de masticación con selección varió ampliamente entre participantes, y con las de otro participante solo hubo selección en E2E3S8 y E2E3S9, sin que ninguna ventana retirara ambas componentes.

5.5.2. Desempeño cuantitativo

En reposo, iCanClean no modificó la señal de E1M1_R2, E1M2_R2 y E2E3S8, porque en ninguna de sus ventanas de reposo se retiró una componente (Tabla 5.5); en ellos, R^2_{reposo} y $Q_{\text{RMS,reposo}}$ coinciden con los de la identidad (Tabla 5.6). Por lo tanto, en esos participantes, el máximo de R^2_{reposo} indica solo que el método no actuó en reposo. En E2E3S7 y E2E3S9, que tuvieron una ventana de reposo con selección, R^2_{reposo} fue 97,2 % y 99,9 %, respectivamente.

		Reposo			Masticación		
			Componentes retiradas			Componentes retiradas	
Participante	Referencias	Ventanas	Una	Ambas	Ventanas	Una	Ambas
E1M1_R2	Propias	4	0	0	2	1	0
	Propias, ventanas de (a)	4	0		2	1	
	Otro participante (a)	4	0	0	2	0	0
E1M2_R2	Propias	6	0	0	28	24	2
	Propias, ventanas de (a)	6	0		26	24	
	Otro participante (a)	6	0	0	26	0	0
E2E3S7	Propias	6	1	0	12	3	0
	Propias, ventanas de (a)	6	1		12	3	
	Otro participante (a)	6	0	0	12	0	0
E2E3S8	Propias	6	0	0	12	8	4
	Propias, ventanas de (a)	6	0		12	12	
	Otro participante (a)	6	1	0	12	1	0
E2E3S9	Propias	6	1	0	12	2	8
	Propias, ventanas de (a)	6	1		12	10	
	Otro participante (a)	6	0	0	12	3	0

Tabla 5.5: iCanClean: número de ventanas de iCanClean de cada condición y número de ellas en que se retiró una componente o ambas, con la configuración elegida para cada participante: con las referencias propias, en todas las ventanas de evaluación y en las que cubre el control (a), donde se indica solo el número de ventanas con alguna componente retirada, y con las referencias de otro participante en esas mismas ventanas. Con ventanas de 4 s, cada ventana de iCanClean abarca un par de ventanas de evaluación, o una sola si es la última de una zona con un número impar de ellas (Sección 4.4.4).

	Reposo				Masticación				Puntaje	
Participante	R^2	CC	Q_{RMS}	E_{rel}	ΔRMS	Reducción en 20–45 Hz	CC	E_{rel}	S_{bal}	$S_{\mathrm{bal,banda}}$
E1M1_R2	100,0 %	1,00	1,00	0	18,3 %	0,79 %	0,817	0,536	59,1	50,4
E1M2_R2	100,0 %	1,00	1,00	0	5,57 %	−40,8 %	0,944	0,324	52,8	29,6
E2E3S7	97,2 %	0,986	0,986	0,158	1,38 %	2,83 %	0,986	0,149	49,3	50,0
E2E3S8	100,0 %	1,00	1,00	0	17,4 %	7,40 %	0,826	0,523	58,7	53,7
E2E3S9	99,9 %	0,999	0,999	0,0260	10,9 %	12,2 %	0,891	0,412	55,4	56,0
Media	99,4 %	0,997	0,997	0,0368	10,7 %	−3,51 %	0,893	0,389	55,1	48,0
Intervalo bootstrap 95 %, límite inferior	98,3 %	0,991	0,991	0	4,97 %	−22,4 %	0,835	0,261	51,9	38,6
Intervalo bootstrap 95 %, límite superior	100,0 %	1,00	1,00	0,100	16,5 %	8,93 %	0,950	0,506	58,2	54,4

Tabla 5.6: iCanClean: métricas de cada participante, con la configuración elegida con los otros cuatro participantes, su media y los límites del intervalo bootstrap, como en la Tabla 5.2.

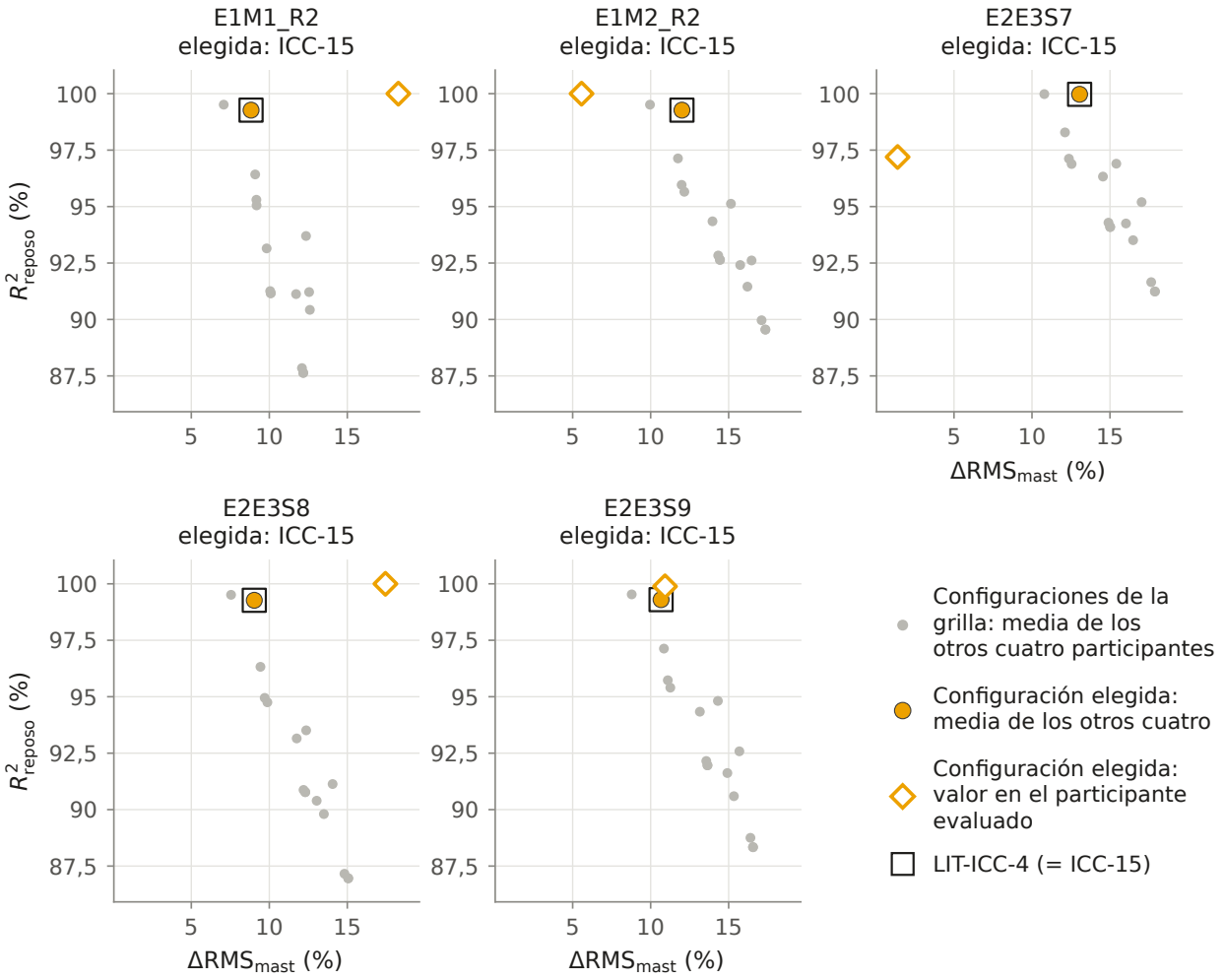


Figura 5.13: iCanClean: relación entre R^2_{reposo} y $\Delta\text{RMS}_{\text{mast}}$ de todas las configuraciones de la grilla, como en la Figura 5.1. El cuadrado abierto ubica LIT-ICC-4, que coincide con la configuración elegida (Tabla B.4).

En masticación, $\Delta\text{RMS}_{\text{mast}}$ varió ampliamente entre participantes, desde 1,38 % en E2E3S7 hasta 18,3 % en E1M1_R2, con una media de 10,7 %. La magnitud de la reducción no siguió el número de ventanas con selección: en E1M2_R2 se retiraron componentes en casi todas las ventanas de masticación, y la reducción fue menor que en E1M1_R2, donde proviene de una sola ventana de iCanClean.

La reducción media de potencia en 20–45 Hz fue negativa, −3,51 %, por el valor de E1M2_R2, −40,8 %: en ese participante, la potencia de esa banda aumentó después de aplicar iCanClean, como la de beta y gamma (Tabla C.2), aunque la RMS en masticación disminuyó en 5,57 %.

La salida de iCanClean presenta, además, discontinuidades en los bordes entre ventanas consecutivas: el salto máximo de la estimación del artefacto, en un borde y un canal primario, fue 1,21 en E1M2_R2, es decir, mayor que la RMS de la entrada en esas ventanas (Tabla C.1).

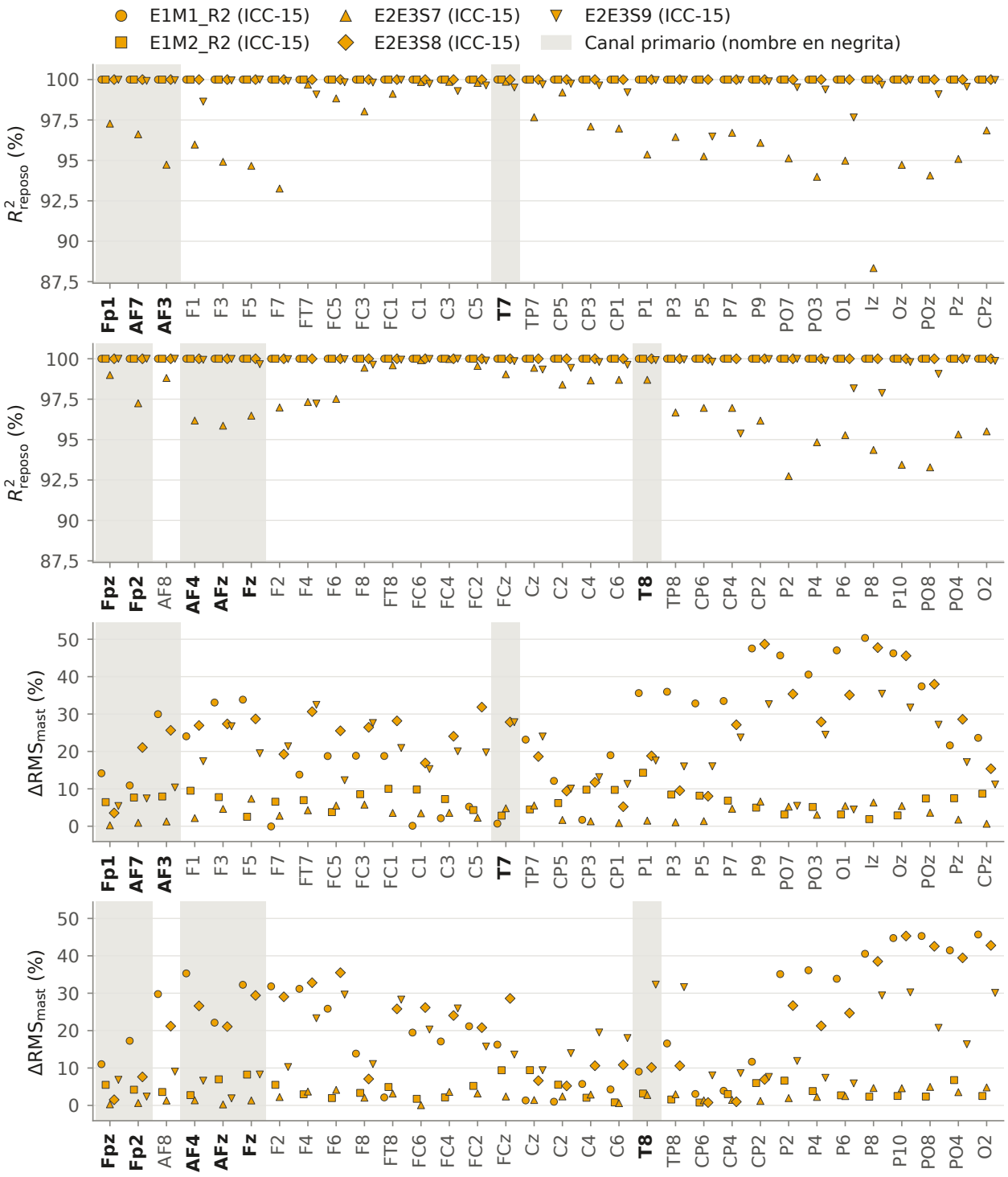


Figura 5.14: iCanClean: R^2_{reposo} (paneles superiores) y $\Delta\text{RMS}_{\text{mast}}$ (paneles inferiores) en cada canal EEG, con la configuración elegida para cada participante, como en la Figura 5.2.

En la Figura 5.14, R^2_{reposo} no cambió en ningún canal de E1M1_R2, E1M2_R2 y E2E3S8, y disminuyó en casi todos los canales de E2E3S7. En masticación, la reducción por canal fue más baja y menos variable en E1M2_R2 y E2E3S7, y en los demás participantes alcanzó sus mayores valores en canales posteriores, fuera de los primarios.

5.5.3. Inspección visual en AF3 y AF4

En reposo, iCanClean no retiró ninguna componente en E1M1_R2, y la señal corregida coincide con la entrada en AF3 y AF4 (Figuras 5.15 y 5.16).

En masticación, solo la ventana de iCanClean que abarca la primera ventana de evaluación tuvo una componente retirada. En su traza, la salida atenúa en ambos canales las deflexiones lentas de gran amplitud y conserva las fluctuaciones más rápidas. La PSD de la salida, que promedia todas las ventanas de masticación, es menor en la parte más baja de delta y, en 20–45 Hz, menor en buena parte de la banda en AF4, pero similar o mayor en AF3. El espectrograma cubre la otra ventana de iCanClean, sin componentes retiradas, por lo que la entrada y la salida no se distinguen.

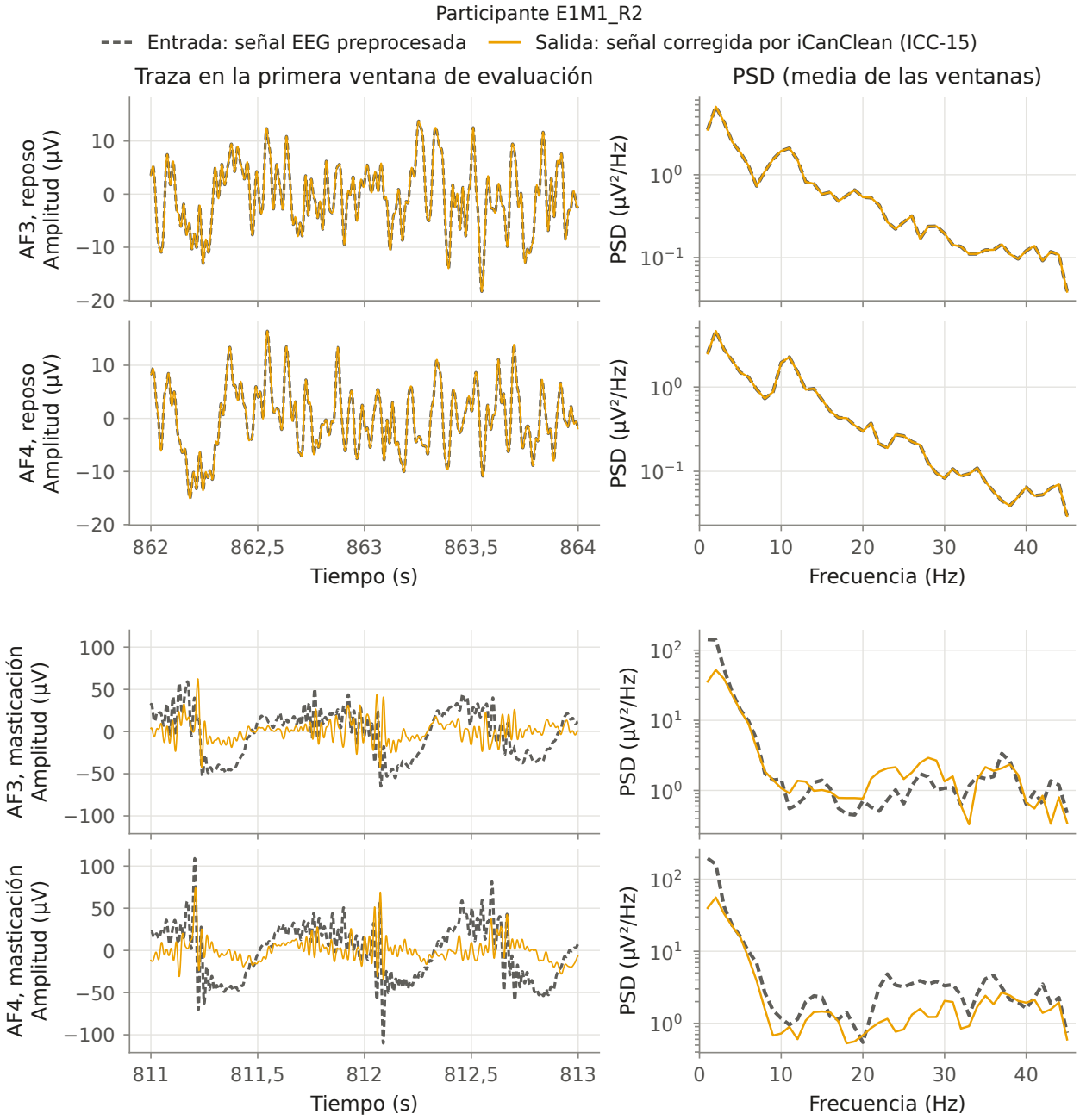


Figura 5.15: iCanClean: inspección visual en AF3 y AF4, participante E1M1_R2, con la configuración elegida para ese participante (Tabla B.1). Filas, paneles y trazos como en la Figura 5.3.

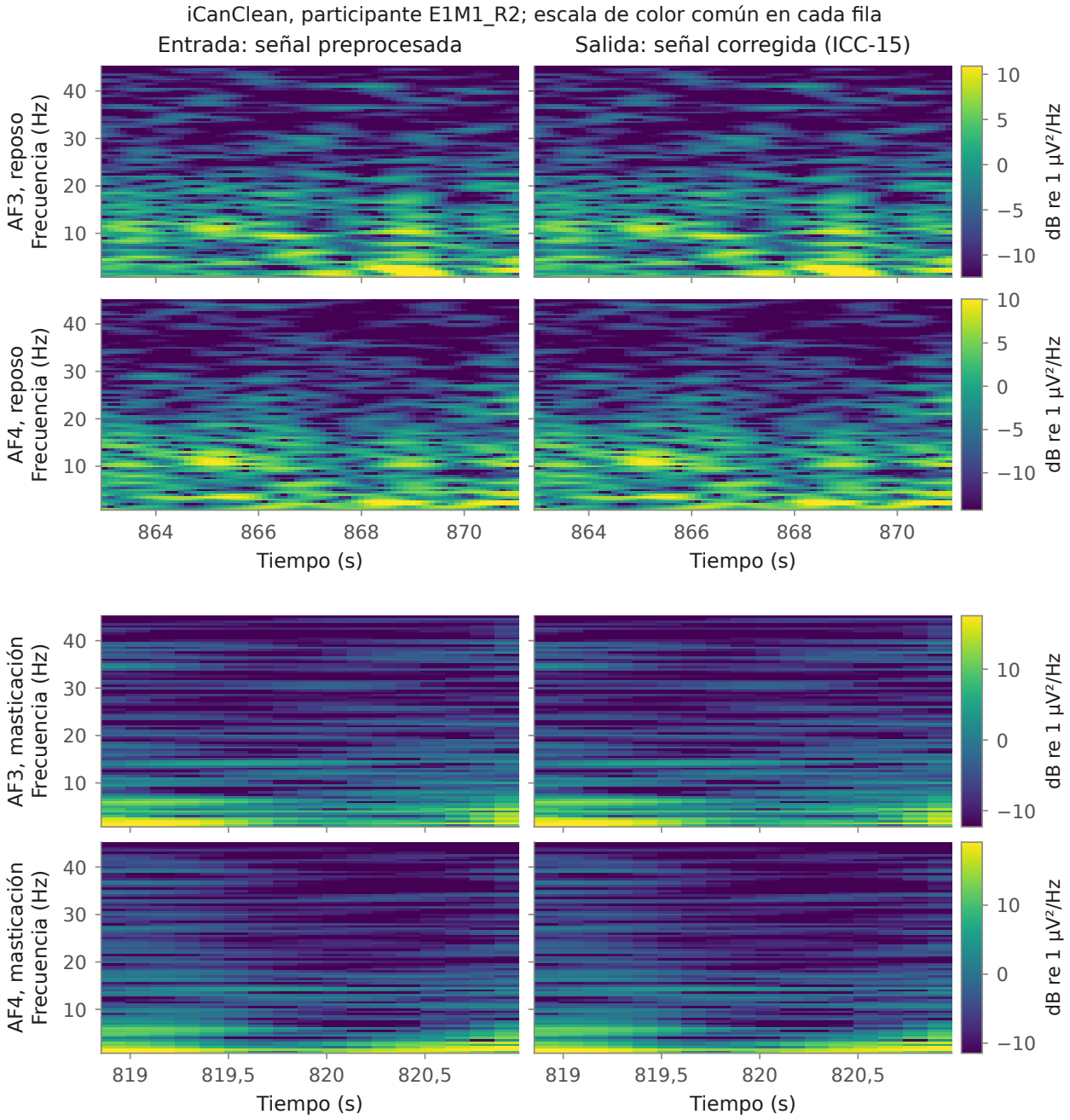


Figura 5.16: iCanClean: espectrogramas de la entrada (izquierda) y de la salida (derecha) en AF3 y AF4, participante E1M1_R2, en reposo y en masticación, con la misma configuración que la Figura 5.15. Tramo y escala de color como en la Figura 5.4.

5.6. Comparación de métodos

5.6.1. Comparación global

La Tabla 5.7 muestra que los métodos se distribuyen de manera distinta entre la reducción de amplitud en masticación y la modificación del reposo. ASR ocupó un extremo, con la mayor reducción media de RMS en masticación, 60,6 %, y la menor similitud en reposo, R^2_{reposo} de 61,1 %, junto con una atenuación del reposo ($Q_{\text{RMS,reposo}}$ de 0,747). El AF por bloques, la CCA con rezagos e iCanClean obtuvieron medias de S_{bal} próximas entre sí, de 55,9, 54,7 y 55,1, con diferencias menores que la variación de cada método entre participantes. Por lo tanto, con cinco participantes, estas diferencias se describen, pero no permiten ordenar los métodos.

	Reposo		Masticación		Puntaje	
Método	R^2	Q_{RMS}	ΔRMS	Reducción en 20–45 Hz	S_{bal}	$S_{\text{bal,banda}}$
Identidad	100,0 %	1,00	0 %	0 %	50,0	50,0
ASR	61,1 %	0,747	60,6 %	65,9 %	60,8	63,5
AF por bloques	97,7 %	0,981	14,1 %	1,13 %	55,9	49,4
CCA con rezagos	98,9 %	0,992	10,6 %	4,18 %	54,7	51,5
iCanClean	99,4 %	0,997	10,7 %	−3,51 %	55,1	48,0

Tabla 5.7: Media de los participantes de las métricas principales y de los puntajes de cada método, con la configuración elegida para cada participante, y de la identidad. Los valores de cada participante se presentan en las Tablas 5.2–5.6, y las medias de las configuraciones informadas sin selección, en la Tabla B.4.

Frente a la identidad, todos los métodos obtuvieron un S_{bal} mayor en todos los participantes, salvo la CCA con rezagos e iCanClean en E2E3S7. Esta comparación se informa sin valor de mérito, porque S_{bal} no penaliza los cambios de amplitud en reposo (Sección 4.5.6): ASR superó a la identidad en todos los participantes con valores de $Q_{\text{RMS,reposo}}$ desde 0,452, mientras que en el AF, la CCA e iCanClean ese valor fue de al menos 0,960. Por lo tanto, un S_{bal} mayor que el de la identidad no indica por sí solo una limpieza selectiva del artefacto.

Con $S_{\text{bal,banda}}$, que usa la reducción de potencia en 20–45 Hz, la media de la CCA con rezagos quedó sobre el valor de la identidad, mientras que en el AF por bloques y en iCanClean la reducción en esa banda, pequeña o negativa, no compensó la pérdida de similitud en reposo. Como ambos puntajes no están en la misma escala, sus valores no se comparan entre sí, pero el cambio de posición indica que la comparación entre el AF, la CCA e iCanClean depende de la medida de atenuación.

Entre las configuraciones informadas sin selección (Tabla B.4), LIT-ICC-2 obtuvo valores medios próximos a los de la elegida de iCanClean.

5.6.2. Análisis por banda

En masticación (Figura 5.17; Tabla C.2), ASR redujo la potencia de todas las bandas, con razones medias entre 0,178 en delta y 0,448 en alfa. El AF por bloques, la CCA con rezagos e iCanClean,

en cambio, concentraron la atenuación en 1–8 Hz, con razones medias de 0,674, 0,790 y 0,790, respectivamente, frente a 0,199 con ASR, y su reducción media en 20–45 Hz fue pequeña o negativa. Como la reducción de RMS resume toda la banda de 1–45 Hz, en estos métodos puede quedar dominada por esas componentes lentas, que en la región frontal pueden seguir también a la actividad ocular.

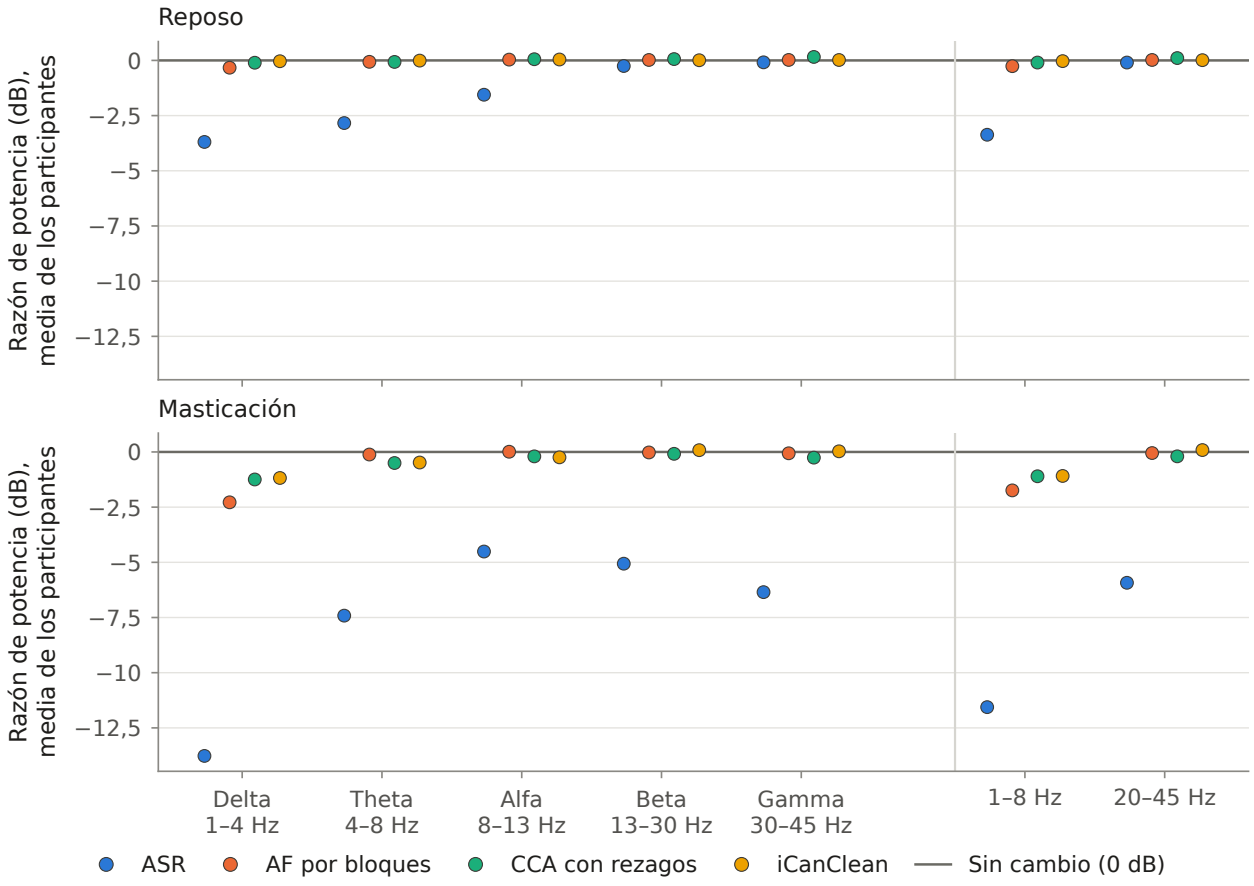


Figura 5.17: Razón de potencia $P_{\text{out}}/P_{\text{in}}$ por banda, en decibeles, en reposo (arriba) y en masticación (abajo), con la configuración elegida para cada participante: media de los participantes de la razón en decibeles de cada uno. Además de las cinco bandas de la Tabla C.2, se incluyen las de 1–8 Hz y de 20–45 Hz, que se solapan con ellas y se separan con una línea vertical. Los métodos se desplazan levemente dentro de cada banda, en el orden de la leyenda. La línea horizontal corresponde a una potencia sin cambio.

En reposo (Tabla C.2), el AF por bloques, la CCA con rezagos e iCanClean mantuvieron las razones de potencia cerca de uno, con razones medias en 1–8 Hz de 0,942, 0,978 y 0,993, mientras que ASR modificó el reposo sobre todo en las bandas bajas y en alfa, con 0,575 en 1–8 Hz. Como la figura promedia la razón en decibeles de cada participante, da más peso que la razón media de las tablas a las atenuaciones más fuertes, lo que se nota cuando la razón varía mucho entre participantes, como en ASR.

5.6.3. Controles de especificidad y de actividad ocular

Con las referencias de otro participante (Sección 4.5.7), la atenuación en masticación disminuyó en el AF por bloques, la CCA con rezagos e iCanClean, pero no en la misma medida (Tablas 5.8 y C.3). En el AF, el control (a) conservó una parte de la atenuación: la fracción $\Delta\mathrm{RMS}_{\mathrm{mast}}(\mathrm{a})/\ \Delta\mathrm{RMS}_{\mathrm{mast}}(\mathrm{propias})$ tuvo una media de 0,282, y de 0,288 con la nueva calibración (a′). En delta, la potencia cambió en la media $-2{,}29$ dB con las referencias propias y $-0{,}61$ dB y $-0{,}59$ dB con (a) y (a′); en reposo, el control (a) reprodujo una parte mayor de ese cambio, $-0{,}20$ dB, frente a $-0{,}33$ dB con las propias.

		$\Delta\mathrm{RMS}_{\mathrm{mast}}$	Fracción respecto de las propias	
Método	Participante	Propias	(a)	(a′)
AF por bloques	E1M1_R2	19,0 %	0,218	0,209
	E1M2_R2	18,9 %	0,305	0,322
	E2E3S7	4,78 %	0,252	0,295
	E2E3S8	14,7 %	0,349	0,328
	E2E3S9	13,3 %	0,288	0,284
	Media	14,1 %	0,282	0,288
CCA con rezagos	E1M1_R2	19,6 %	0,0023	0,0007
	E1M2_R2	3,02 %	0,0401	0,0352
	E2E3S7	2,94 %	0,0380	0,0746
	E2E3S8	17,4 %	0,0067	0,0068
	E2E3S9	10,1 %	0,0151	0,0172
	Media	10,6 %	0,0204	0,0269
iCanClean	E1M1_R2	18,3 %	0	—
	E1M2_R2	5,69 %	0	—
	E2E3S7	1,38 %	0	—
	E2E3S8	17,4 %	0,0101	—
	E2E3S9	10,9 %	0,128	—
	Media	10,7 %	0,0275	—

Tabla 5.8: Fracción de la reducción de RMS en masticación que también se obtiene con las referencias de otro participante, con la configuración elegida para cada participante. $\Delta\mathrm{RMS}_{\mathrm{mast}}$ con las referencias propias y las fracciones con (a) y (a′) se calculan en las ventanas que cubre el control (a), que solo en E1M2_R2 no son todas las de evaluación. En la fila «Media», la fracción es la media de las fracciones de cada participante. En iCanClean, (a′) no se aplica (—).

En la CCA con rezagos, en cambio, la atenuación casi desapareció con esas referencias, con medias de las fracciones de 0,0204 con (a) y 0,0269 con (a′). En iCanClean, el control (a) no produjo atenuación en E1M1_R2, E1M2_R2 y E2E3S7, donde el método no retiró ninguna componente con esas referencias (Tabla 5.5), y la media de las fracciones fue 0,0275. Por lo tanto, en la CCA y en iCanClean la atenuación dependió casi por completo de las referencias propias, y en el AF, solo en parte. Este control indica de qué referencias depende la atenuación, pero no establece que lo retirado sea solo actividad muscular.

Con referencias nulas (b), iCanClean devolvió la entrada exacta en 13 de las 13 piezas de evaluación, y la CCA con rezagos, salvo por el redondeo, con una diferencia relativa máxima de $2{,}38 \cdot 10^{-16}$; en ambos métodos, el control verifica la implementación. El AF por bloques no devolvió la entrada en ninguna pieza (0 de 13; diferencia relativa máxima de 0,144), por el término de sesgo, que incluyen todas sus configuraciones. Ese efecto fue pequeño: en la media, la potencia de delta en masticación cambió $-0{,}04$ dB, frente a $-2{,}28$ dB con las propias en todas las ventanas. Que el AF sin ese término devuelva la entrada se comprobó solo con señales sintéticas.

En masticación (Tabla 5.9), la entrada de 1–8 Hz compartió con el EOG vertical una fracción de su varianza muy distinta entre participantes, de 0,0836 en E2E3S8 a 0,764 en E2E3S7. Salvo en E2E3S7, ASR retiró prácticamente toda la varianza de esa banda, y lo retirado compartió con el EOG una fracción similar a la de la entrada. Lo retirado por el AF, la CCA e iCanClean compartió con el EOG una fracción menor que la entrada en todos los participantes y, salvo en el AF en E2E3S7, una mayor con las referencias EMG, una asociación en parte esperable, porque estos métodos lo estiman a partir de ellas. En reposo, la entrada compartió con el EOG vertical una fracción mayor que en masticación en todos los participantes salvo E1M1_R2. ASR retiró en E1M2_R2, E2E3S8 y E2E3S9 la mayor parte de la varianza de 1–8 Hz, compartida en gran parte con el EOG, mientras que el AF, la CCA e iCanClean retiraron como máximo 0,0660 de esa varianza, el AF en E2E3S7. En el reposo de E1M1_R2, E1M2_R2 y E2E3S8, en que iCanClean no retiró componentes, sus fracciones explicadas no están definidas.

Con la derivación horizontal EXG3−EXG4, usada como sensibilidad, la lectura cambia. Lo retirado por iCanClean compartió con ella una fracción mayor que la entrada en todos los participantes, por ejemplo, 0,353 frente a 0,205 en E2E3S8, y lo mismo ocurrió con la CCA, salvo en E1M2_R2, y con el AF en E2E3S8 y E2E3S9. Esas fracciones fueron pequeñas salvo en E1M2_R2 y E2E3S8, donde llegaron a 0,368 y donde la entrada también compartió más varianza con ella que con la vertical (en E1M2_R2, 0,213). Con los canales EOG monopolares, las fracciones fueron altas con todos los métodos en E1M1_R2 y E2E3S8, de 0,593 a 0,914, pero, como esos canales comparten actividad con el EMG en masticación, no separan la actividad ocular de la muscular.

5.6.4. Criterios para examinar la hipótesis

Las Tablas 5.10 y 5.11 resumen la evaluación de los criterios. En el criterio 1, iCanClean no superó a ASR en ningún participante; frente al AF por bloques, el criterio se cumplió solo en E2E3S8, y frente a la CCA con rezagos, en ningún participante. En E1M2_R2, las diferencias de la reducción en 20–45 Hz fueron las más negativas frente a los tres comparadores, porque en la salida de iCanClean la potencia de esa banda aumentó, y dependen en buena parte de los canales marcados por su gran amplitud, aunque sin ellos no desaparecen (Apéndice D.3).

En el criterio 2, iCanClean modificó el reposo menos que ASR salvo en E2E3S7, donde ASR prácticamente no lo modificó; frente al AF por bloques, todas las comparaciones terminaron en empate, y frente a la CCA con rezagos, el criterio se cumplió solo en E1M2_R2. La ventaja frente

		Masticación			Reposo		
		Explicada por			Explicada por		
Señal	Participante	EOG	EMG	Retirada	EOG	EMG	Retirada
Entrada	E1M1_R2	0,493			0,299		
	E1M2_R2	0,137			0,936		
	E2E3S7	0,764			0,907		
	E2E3S8	0,0836			0,662		
	E2E3S9	0,249			0,909		
Retirada por ASR	E1M1_R2	0,526	0,376	1,01	0,0140	0,0110	0,0401
	E1M2_R2	0,138	0,108	1,00	0,974	0,0396	0,978
	E2E3S7	0,0511	0,0535	0,0614	—	—	0,0000
	E2E3S8	0,0985	0,511	0,954	0,819	0,0125	0,792
	E2E3S9	0,255	0,134	0,953	0,980	0,0096	0,938
Retirada por el AF por bloques	E1M1_R2	0,0128	0,784	0,121	0,0018	0,104	0,0157
	E1M2_R2	0,0003	0,312	0,228	0,0703	0,173	0,0247
	E2E3S7	0,0984	0,0562	0,0873	0,0826	0,269	0,0660
	E2E3S8	0,0216	0,643	0,258	0,0231	0,226	0,0107
	E2E3S9	0,0123	0,233	0,156	0,0216	0,277	0,0338
Retirada por la CCA con rezagos	E1M1_R2	0,0014	0,999	0,291	0,0162	0,979	0,0004
	E1M2_R2	0,0018	0,769	0,0260	0,0361	0,992	0,0219
	E2E3S7	0,0058	0,999	0,0321	0,0548	0,998	0,0387
	E2E3S8	0,0277	0,995	0,317	0,0077	1,00	0,0002
	E2E3S9	0,0108	0,993	0,0903	0,0014	0,999	0,0011
Retirada por iCanClean	E1M1_R2	0,0001	0,921	0,307	—	—	0
	E1M2_R2	0,0023	0,620	0,107	—	—	0
	E2E3S7	0,0144	0,187	0,0102	0,0081	0,280	0,0402
	E2E3S8	0,0163	0,886	0,287	—	—	0
	E2E3S9	0,0106	0,619	0,0824	0,0004	0,200	0,0001

Tabla 5.9: Control de actividad ocular en 1–8 Hz, con la configuración elegida para cada participante: mediana, en los canales frontales, de las cuatro magnitudes de la Sección 4.5.7, como proporción. EOG: fracción de la varianza de la entrada, $x[n]$, o de la señal retirada, $x[n] - y[n]$, que explica la derivación vertical EXG1−EXG2; EMG: la que explican las referencias EMG principales; Retirada: $\|x - y\|^2/\|x\|^2$, que no está acotada por uno. Las fracciones explicadas no están definidas cuando el método no retira nada (—). En el reposo de E2E3S7, ASR no modificó la señal, y esas fracciones, que se calcularían sobre un residuo de redondeo, no se informan.

a ASR se lee junto con la actividad de iCanClean en reposo: en E1M1_R2, E1M2_R2 y E2E3S8 no retiró ninguna componente en esa condición (Tabla 5.5), de modo que en ellos su ventaja indica que dejó sin cambios el EEG observado, y no una mayor preservación de la actividad cerebral. Además, en E1M2_R2, E2E3S8 y E2E3S9, la modificación del reposo de ASR incluyó actividad frontal de 1–8 Hz compartida en gran parte con el EOG vertical, que iCanClean prácticamente no retiró.

El requisito previo se cumplió en todos los participantes, aunque es poco exigente para iCanClean (Sección 4.5.8). En conjunto, frente a ASR el criterio 1 no se cumple y el criterio 2 se cumple, y frente al AF por bloques y a la CCA con rezagos no se cumple ninguno de los dos criterios. Como el criterio 1 no se cumple frente a ningún comparador, la etiqueta de la hipótesis es «no respaldada», y

	Criterio 1: atenuación en masticación			Criterio 2: modificación en reposo		
Comparador	Cumple	No evaluables	Veredicto	Cumple	No evaluables	Veredicto
ASR	0	0	no se cumple	4	0	se cumple
AF por bloques	1	0	no se cumple	0	0	no se cumple
CCA con rezagos	0	0	no se cumple	1	0	no se cumple
Identidad (requisito previo)	5	—	se cumple		—	
Etiqueta de la hipótesis: no respaldada						

Tabla 5.10: Evaluación descriptiva de los criterios de la Sección 4.5.8 para iCanClean, con la configuración elegida para cada participante, frente a cada comparador y, como requisito previo, frente a la identidad: participantes, de cinco, en que se cumple cada criterio, comparaciones no evaluables y veredicto. El requisito previo considera solo la atenuación en masticación, y la identidad tiene valor en todos los participantes (—).

		Criterio 1: atenuación en masticación			Criterio 2: modificación en reposo			
Comparador	Participante	$\Delta\mathrm{RMS}$	20–45 Hz	Veredicto	R^2	$\lvert 1-Q_{\mathrm{RMS}}\rvert$	Banda (dB)	Veredicto
Margen		1	1		1	0,01	0,5	
ASR	E1M1_R2	−47,2	−65,0	no cumple (estable)	7,48	0,0360	1,25	cumple (frágil)
	E1M2_R2	−80,2	−128	no cumple (estable)	76,8	0,548	9,35	cumple (estable)
	E2E3S7	−8,19	−7,20	no cumple (estable)	−2,81	−0,0142	−0,20	no cumple (estable)
	E2E3S8	−53,9	−77,8	no cumple (estable)	51,0	0,270	4,81	cumple (estable)
	E2E3S9	−59,9	−69,8	no cumple (estable)	59,2	0,413	3,92	cumple (estable)
AF por bloques	E1M1_R2	−0,74	−0,98	empate	0,98	0,0091	0,21	empate
	E1M2_R2	−13,0	−41,8	no cumple	2,05	0,0229	0,36	empate
	E2E3S7	−3,39	2,34	no cumple	1,26	0,0262	0,45	empate
	E2E3S8	2,73	6,90	cumple	0,60	0,0091	0,20	empate
	E2E3S9	−2,39	10,3	no cumple	3,69	0,0140	0,22	empate
CCA con rezagos	E1M1_R2	−1,30	−6,02	no cumple	0,20	0,0016	0,02	empate
	E1M2_R2	2,67	−32,0	no cumple	2,10	0,0104	0,65	cumple
	E2E3S7	−1,55	−2,04	no cumple	0,43	0,0134	0,14	empate
	E2E3S8	0,07	4,11	empate	0,01	0,0007	0,01	empate
	E2E3S9	0,79	−2,50	no cumple	0,02	0,0005	0,03	empate
Identidad	E1M1_R2	18,3		cumple				
	E1M2_R2	5,57		cumple				
	E2E3S7	1,38		cumple				
	E2E3S8	17,4		cumple				
	E2E3S9	10,9		cumple				

Tabla 5.11: Matriz de los criterios de la Sección 4.5.8 para iCanClean, con la configuración elegida para cada participante: diferencia de cada métrica con cada comparador y resultado de cada comparación; al final, requisito previo frente a la identidad. Las diferencias de $\Delta\mathrm{RMS}_{\mathrm{mast}}$, de la reducción de potencia en 20–45 Hz (en puntos porcentuales) y de R^2_{reposo} (en puntos) son de iCanClean menos el comparador, y las de $\lvert 1-Q_{\mathrm{RMS,reposo}}\rvert$ y de la mayor desviación absoluta de la razón de potencia de las cinco bandas (en dB), del comparador menos iCanClean, de modo que una diferencia positiva favorece siempre a iCanClean. La primera fila indica el margen de cada métrica, con el que se comparan las diferencias sin redondear. Frente a ASR se indica, entre paréntesis, si la comparación es estable o frágil ante las repeticiones de ASR.

su rango, sin criterios indeterminados, «no respaldada».

Frente a ASR, cada comparación se recalculó con las 6 repeticiones de sensibilidad de ASR, en las que el número de calibraciones fallidas fue 0. De las 10 comparaciones, 9 fueron estables y 1 fue frágil, la del criterio 2 en E1M1_R2: con la calibración desplazada dos y ocho muestras y con O1 retirado, el R^2_{reposo} de ASR fue 99,7 %, 100,0 % y 100,0 %, tan cercano al de iCanClean, 100,0 %, que la comparación dejó de cumplirse y, con ella, el criterio 2 frente a ASR en esas repeticiones (Apéndice D.1). La etiqueta no cambió (el número de repeticiones con otra etiqueta fue 0), porque el criterio 1 no se cumple frente a ningún comparador.

Como los criterios no incluyen la especificidad de la atenuación, la etiqueta se lee junto con los controles de la Sección 5.6.3. La media de las fracciones $\Delta\mathrm{RMS}_{\mathrm{mast}}(\mathrm{a})/\Delta\mathrm{RMS}_{\mathrm{mast}}(\mathrm{propias})$ de iCanClean, 0,0275, indica que su atenuación dependió de sus propias referencias, como la de la CCA con rezagos, algo que el requisito previo no permite distinguir. Lo que iCanClean retiró en masticación compartió con el EOG vertical menos varianza en 1–8 Hz que la entrada en todos los participantes, como en el AF y la CCA, y con la derivación horizontal, más. La etiqueta resume los criterios y no califica la calidad de la limpieza de iCanClean ni la de los comparadores.

Como las repeticiones de ASR, las sensibilidades de diseño que recalculan la etiqueta no la cambiaron (Apéndice D). Con los canales monopolares filtrados en 1–45 Hz como referencia, iCanClean atenuó más en masticación que con la referencia principal, pero también modificó más el reposo, y con las referencias filtradas en 20–120 Hz no modificó la señal. Con ventanas de 2 s, retiró componentes también con las referencias de otro participante, en reposo con una frecuencia igual o mayor que con las propias en E1M1_R2, E1M2_R2, E2E3S8 y E2E3S9, y ninguna duración de la ventana obtuvo el mayor S_{bal} en todos los participantes.

5.7. Evaluación semisintética

La Tabla C.4 presenta las métricas frente a $\boldsymbol{X}_{\mathrm{rep}}$ de cada participante y su media, en cada nivel de SNR. La identidad no modifica la mezcla, de modo que sus métricas describen la contaminación agregada: un método acerca su salida a $\boldsymbol{X}_{\mathrm{rep}}$ cuando obtiene un error menor o un CC mayor que ella. El AF por bloques, la CCA con rezagos e iCanClean obtuvieron, en la media y en los tres niveles de SNR, un $\mathrm{RRMSE}_{\mathrm{t}}$ y un $\mathrm{RRMSE}_{\mathrm{f}}$ menores que los de la identidad y un CC mayor, con una diferencia más marcada en el error espectral: con $\mathrm{SNR} = -5$ dB, su $\mathrm{RRMSE}_{\mathrm{f}}$ medio fue 3,57, 4,82 y 5,09, respectivamente, frente a 6,29 en la identidad. ASR mostró un patrón distinto: con $\mathrm{SNR} = -5$ dB y con $\mathrm{SNR} = 0$ dB, su $\mathrm{RRMSE}_{\mathrm{f}}$ medio fue mucho menor que el de la identidad, 1,03 y 0,762, pero su CC medio fue menor que el de la identidad en los tres niveles, y con $\mathrm{SNR} = 5$ dB su $\mathrm{RRMSE}_{\mathrm{t}}$ medio fue mayor. Estas métricas se leen con el alcance de la Sección 4.7.

5.8. Discusión y limitaciones

5.8.1. Compromiso entre la atenuación en masticación y la modificación del reposo

Entre la atenuación en masticación y la modificación del reposo, ASR ocupó un extremo, y el AF por bloques, la CCA con rezagos e iCanClean formaron un grupo de atenuación moderada y poca modificación del reposo, que los resultados no permiten ordenar (Sección 5.6.1). Esta distribución describe a cada método con la configuración que eligió el puntaje balanceado, y no su comportamiento en toda la grilla: en todos los métodos, esa configuración quedó en el extremo de menor modificación media del reposo de su grilla o cerca de él, en ASR y en la CCA con rezagos a menudo por el desempate, que favorece esa condición (Tabla B.2).

En ASR, la atenuación en masticación correspondió más a una reconstrucción de la señal que a la resta de una componente. En E1M2_R2, E2E3S8 y E2E3S9 retiró, además, la mayor parte de la varianza frontal de 1–8 Hz del reposo, la condición con menor contaminación muscular, de modo que su atenuación no se restringió al artefacto; lo retirado puede corresponder en buena parte a actividad compartida con el EOG vertical, aunque el control no descarta que incluya actividad cerebral (Secciones 5.6.3 y 3.2).

En el AF por bloques, la atenuación se concentró en la banda delta, y una parte apreciable de ella se obtuvo también con las referencias de otro participante, mientras que con referencias nulas, con solo el término de sesgo, fue pequeña (Sección 5.6.3). Por lo tanto, dependió en parte de las referencias propias, pero no puede atribuirse por completo al artefacto.

En la CCA con rezagos, la atenuación casi desapareció con las referencias de otro participante, de modo que dependió de las referencias propias, aunque su magnitud varió mucho entre participantes. Sin embargo, como sus umbrales quedaron cerca del nivel de azar en las piezas cortas de los datos de desarrollo (Apéndice A.3), el umbral no basta para atribuir al artefacto los pares retirados.

En iCanClean, la poca modificación del reposo se explica porque, con ventanas de 4 s y el mayor umbral de la grilla, el método rara vez retiró componentes en esa condición (Sección 5.5.2). Su atenuación fue moderada y se concentró en las bandas bajas, y su reducción media en 20–45 Hz fue negativa por el aumento de esa banda en E1M2_R2. Una interpretación posible, que no se examinó, es que la proyección, estimada por mínimos cuadrados sobre toda la banda, quede dominada por las componentes de baja frecuencia, de mayor energía, y que la combinación de canales que se resta agregue a cada canal actividad de 20–45 Hz de otros canales. Su resultado depende, además, de la duración de la ventana: con 2 s, atenuó más en masticación, pero modificó más el reposo y seleccionó componentes con frecuencia también con las referencias de otro participante (Apéndice D.4).

Por lo tanto, en los métodos con referencia, la reducción de RMS en masticación reflejó sobre todo una atenuación de 1–8 Hz, cuyo origen muscular no está establecido, y no de 20–45 Hz (Sección 5.6.2). Esto es relevante para los estudios que examinan la actividad θ durante la masticación [7], que motivan esta tesis: la CCA con rezagos e iCanClean redujeron la potencia de θ en masticación (Tabla C.2), sin que esta evaluación permita establecer si lo retirado era artefacto o

actividad cerebral.

5.8.2. Lectura de la etiqueta de la hipótesis

La etiqueta de la hipótesis fue «no respaldada» porque el criterio de atenuación en masticación no se cumplió frente a ningún comparador (Sección 5.6.4), y, como los criterios no incluyen la especificidad de la atenuación, se lee junto con los controles y no como una calificación de la limpieza. La dependencia de las referencias propias y la menor varianza compartida con el EOG vertical no distinguen a iCanClean de la CCA con rezagos ni, en el caso del EOG, del AF por bloques; con la derivación horizontal, en cambio, lo retirado por iCanClean compartió más varianza que la entrada (Sección 5.6.3).

La etiqueta tiene los sesgos descritos en la Sección 4.5.8. El sesgo a favor de iCanClean tuvo poco espacio con la configuración elegida, porque su umbral superó el percentil 95 del nivel de azar en reposo y prácticamente coincidió con él en masticación. En contra de iCanClean, el criterio de atenuación frente a ASR no se cumplió. Sin embargo, la etiqueta no se debe solo al sesgo en contra: aun sin considerar ASR, el criterio de atenuación tampoco se cumplió frente al AF por bloques ni frente a la CCA con rezagos.

5.8.3. Coherencia entre las métricas, la prueba semisintética y la inspección visual

Las métricas fueron coherentes entre sí y se leen en conjunto, y la prueba semisintética coincidió en dirección con ellas (Sección 5.7). Los métodos con referencia acercaron su salida a $\boldsymbol{X}_{\mathrm{rep}}$ más que la identidad, aunque de forma moderada en el dominio temporal, y en ASR, el menor error espectral con las SNR más bajas, junto con una correlación menor que la de la identidad, concuerda con una salida que se acerca al espectro del objetivo sin conservar su forma. Estas coincidencias se leen con el alcance de la prueba, que favorece a los métodos con referencia; en E1M1_R2, además, sus mezclas son breves y comparten el tramo de reposo. Que el acercamiento temporal fuera moderado es esperable en parte, porque la contaminación agregada incluye actividad cerebral ausente en $\boldsymbol{X}_{\mathrm{rep}}$, y es compatible con que parte de la contaminación no se relacione linealmente con las referencias, algo que la prueba no permite separar.

La inspección visual de E1M1_R2 concordó con sus métricas en todos los métodos, pero en ese participante la atenuación de los métodos con referencia provino sobre todo de la primera ventana de masticación, contigua a la reserva: sin ella, la $\Delta\mathrm{RMS}_{\mathrm{mast}}$ de iCanClean fue 0 %, las del AF por bloques y de la CCA con rezagos fueron mucho menores, y la reducción en 20–45 Hz de ASR pasó de 65,8 % a 4,17 % (Apéndice D.3). Esa ventana pertenece, además, a la única pieza en que la continua residual superó el umbral de verificación, y como toda la atenuación de iCanClean, que no se calibra, provino también de ella, la cercanía a la calibración no basta para explicarla. Quedan lecturas distintas, que la evaluación no permite distinguir: que el artefacto evaluado en E1M1_R2 se concentre en esa ventana, lo que concuerda con que la reducción en 20–45 Hz de ASR, el indicador más cercano a la actividad muscular, provenga casi por completo de ella, o que la atenuación de baja

frecuencia de los métodos con referencia refleje en parte componentes lentas propias de esa pieza.

5.8.4. Relación con la literatura

La relación de estos resultados con la literatura depende de las diferencias de diseño. En un fantoma con fuentes cerebrales conocidas, iCanClean obtuvo el puntaje más alto frente a ASR, a una CCA automática y a un filtrado adaptativo [12], con una métrica que requiere conocer esas fuentes, con la configuración que la maximizó en cada condición y, para el músculo facial, con seudorreferencias. Por lo tanto, los resultados de esta tesis, obtenidos con referencias EMG, configuraciones elegidas sin los datos de cada participante y métricas referidas al EEG observado, no confirman ni contradicen ese hallazgo.

La configuración de iCanClean elegida en todos los participantes coincidió con LIT-ICC-4, que Gonsisko et al. obtuvieron con muchos más canales de referencia y con otro criterio de calidad [58]. En ese trabajo, el umbral óptimo disminuyó al reducir el número de canales de referencia; con solo dos referencias, en cambio, la selección retuvo el mayor umbral de la grilla, y los umbrales menores quedaron bajo el percentil 95 del nivel de azar (Sección 5.5.1). Esto sugiere que, con pocas referencias, el umbral debe leerse frente al nivel de azar medido en los propios datos, más que tomarse de otros montajes.

En ASR, la selección eligió las mayores tolerancias de la grilla, superiores a los valores típicos que indican sus autores [46] y dentro de los intervalos óptimos que Chang et al. informaron en conducción simulada [48, 50]. Aun así, ASR modificó de forma marcada el reposo de varios participantes y dependió de detalles de la calibración, lo que concuerda con la dependencia de sus parámetros y de la calidad de la calibración que describe la literatura [48, 49].

En un trabajo aún no revisado por pares, la mayor supresión durante la masticación se obtuvo al combinar iCanClean, con seudorreferencia, y ASR [7]. Esta tesis no evaluó esa combinación, pero sus resultados indican que la supresión no basta para comparar procedimientos de limpieza, porque la mayor atenuación de ASR se acompañó de la mayor modificación del reposo. La heterogeneidad entre participantes, que en varios casos superó a las diferencias entre métodos, concuerda, además, con que la distribución de la contaminación facial sobre el cuero cabelludo podría depender de cada sujeto [30], y con que, frente a contaminación miogénica real, ningún enfoque dominó [45].

5.8.5. Limitaciones

La principal limitación es que la evaluación no dispone de una referencia de la actividad cerebral: todas sus métricas, incluidas las de la prueba semisintética, comparan la salida con un EEG observado, que contiene también actividad no cerebral. Por lo tanto, la similitud con el EEG observado no equivale a preservar la actividad cerebral, y una reducción de amplitud no indica qué parte de lo retirado era artefacto.

Con cinco participantes, las diferencias se describen sin atribuirles significación, y, con la regla de al menos cuatro de cinco participantes, un criterio que se cumple justo en ese mínimo depende de un

solo participante. Es el caso del criterio de modificación del reposo frente a ASR, cuya comparación en E1M1_R2 cambió con algunas repeticiones de ASR, sin que cambiara la etiqueta (Sección 5.6.4).

El costo de calibrar con datos del propio participante recayó sobre todo en E1M1_R2: conserva solo 3 ventanas de masticación, y su reserva de masticación tiene poca actividad EMG en los primeros segundos. Por ello, sus métricas de masticación, incluido el S_{bal} de ASR, el mayor de la evaluación, se leen con especial cautela. Además, cada reserva queda a solo 4 s de la primera ventana evaluada de su pieza, lo que puede favorecer a los métodos calibrados; sin esas ventanas, el S_{bal} medio bajó en todos los métodos, también en iCanClean, que no se calibra, sin cambiar los veredictos agregados ni la etiqueta (Apéndice D.3).

El diseño temporal también limita la comparación entre las condiciones (Sección 4.3): en E2, la diferencia entre ellas incluye los cambios del registro entre el reposo y la masticación y la actividad de la tarea que acompaña a esta, y en E1M1_R2 el reposo podría conservar efectos de la masticación que lo precede.

Los intervalos se delimitaron con el EMG y el EEG, en su mayoría propuestos con asistencia de inteligencia artificial y ratificados visualmente, y no con los códigos de evento, cuyo significado no se conoce (Sección 4.1.5). Esta delimitación favorece tramos de masticación con ráfagas claras y de reposo con poca actividad EMG, de modo que la evaluación no representa necesariamente una masticación menos intensa ni un reposo con más actividad muscular [31].

Los criterios y las demás decisiones de diseño no son independientes de estos registros (Sección 4.5.8). Con las ventanas de masticación independientes del diseño, que son menos y no incluyen E1M1_R2, el criterio de atenuación frente a ASR y frente a la CCA con rezagos dio el mismo resultado que con todas las ventanas, frente al AF por bloques se cumplió, además, en E2E3S9, y el requisito previo quedó en empate en E2E3S7 (Apéndice D.3).

S_{bal} no penaliza los cambios de amplitud del reposo, a diferencia del criterio 2 (Sección 4.5.6): así, ASR superó a la identidad en S_{bal} en todos los participantes aun atenuando el reposo de varios de ellos. El desempate por $|1 - Q_{\mathrm{RMS,reposo}}|$ reduce este desajuste, pero no lo elimina. Además, como las configuraciones elegidas quedaron en un borde de la grilla o junto a él (Tablas B.1 y B.3), una grilla más amplia podría desplazar la posición de cada método en el compromiso.

Las referencias EMG se registraron solo sobre los maseteros, aunque el temporal y el frontal contaminan el EEG incluso en contracciones débiles [17] y en reposo [31], de modo que pueden no representar toda la actividad muscular del EEG. Tampoco se aplicó la recomendación de los autores de iCanClean de rerreferenciarlas a su propio promedio [12]: con los canales monopolares referidos al promedio del EEG, examinados solo como sensibilidad, iCanClean atenuó más en masticación, pero también modificó más el reposo (Apéndice D.2), como advierten esos autores, aunque la evaluación no permite establecer que lo retirado fuera actividad cerebral. Las referencias principales, filtradas en 1–45 Hz, incluyen además componentes cuyo origen puede no ser muscular, mientras que con las de 20–120 Hz iCanClean no modificó la señal, quizás porque estas solo comparten con el EEG la banda de 20–45 Hz, algo que no se examinó.

La actividad ocular puede contribuir a la atenuación en masticación, porque $\Delta\mathrm{RMS}_{\mathrm{mast}}$ resume

la banda de 1–45 Hz y la entrada de los canales frontales compartió con el EOG vertical una fracción muy variable de su varianza en 1–8 Hz (Tabla 5.9). Además, el control no descarta que la atenuación de los métodos con referencia incluya actividad compartida con el EOG horizontal (Sección 5.6.3). Por ello, $\Delta\mathrm{RMS}_{\mathrm{mast}}$ no puede leerse solo como una reducción del artefacto muscular, y la reducción de potencia en 20–45 Hz es, entre los indicadores utilizados, el más cercano a la actividad muscular.

En ASR, la calibración quedó por debajo del minuto de reposo que se recomienda [49], y el descarte de ventanas atípicas la redujo aún más (Tabla 5.1). Con una calibración breve, los umbrales pueden representar mal la variabilidad del reposo, lo que concuerda con la fragilidad de sus resultados ante pequeños cambios de la calibración (Sección 5.2.2).

En iCanClean, cada ventana se limpia con una proyección propia, y la salida presenta discontinuidades entre ventanas consecutivas, cuyo salto máximo, en un canal primario de E1M2_R2, superó la RMS de la entrada (Tabla C.1); las métricas por ventana no las registran, pero pueden introducir componentes de banda ancha en los análisis de la señal continua. Además, con ventanas de 4 s, la última ventana de algunas piezas se estimó en parte con la guarda final, aunque excluirla no cambió los veredictos agregados ni la etiqueta (Apéndice D.3).

En el preprocesamiento común, la envolvente del EMG, parte de la configuración elegida del AF por bloques, puede presentar transitorios no corregidos en la primera y en la última ventana evaluada de cada pieza (Sección 4.3), y la referencia promedio reparte en todos los canales una parte de la actividad de cada uno, sin una regla para retirar canales defectuosos.

Las implementaciones de ASR, que adapta la biblioteca asrpy, y de iCanClean, que sigue el algoritmo publicado con un ventaneo distinto del de la implementación oficial (Sección 4.4.4), están escritas en Python. Su equivalencia numérica con las implementaciones de referencia en MATLAB no se comprobó, de modo que los resultados corresponden a las implementaciones de esta tesis.

Por último, la evaluación se limita a los intervalos anotados (Sección 4.1.5). La búsqueda de masticación con el EMG en registros de E1 no evaluados (Apéndice G) encontró en los archivos `ch` de E1M1_R2 y de E1M2_R2 un patrón compatible con masticación de unos 424,5 s y 148,0 s, respectivamente, de modo que, en E1, la masticación evaluada es solo una parte de la actividad compatible con masticación de esos registros. Esas duraciones describen un patrón EMG y no la conducta.

CAPÍTULO

6

Conclusiones

En esta tesis se evaluó iCanClean, en su formulación publicada, para atenuar los artefactos de masticación en el EEG, y se comparó con ASR, el AF por bloques y la CCA con rezagos mediante un procedimiento común, con una configuración elegida para cada participante sin utilizar sus datos. Este capítulo responde a los objetivos, evalúa la hipótesis y propone recomendaciones.

6.1. Respuesta a los objetivos

En relación con el primer objetivo específico, iCanClean se implementó en el entorno común en Python según su formulación publicada, con las derivaciones bipolares de ambos maseteros como referencias. Con referencias nulas, la implementación devolvió la entrada exacta en 13 de las 13 piezas de evaluación, lo que la verifica en ese caso, aunque su equivalencia numérica con la implementación oficial no se comprobó (Sección 5.8.5). La selección eligió para todos los participantes la misma configuración, que coincide con la de la literatura con ventanas de 4 s.

En cuanto al segundo objetivo, la variante extendida propia, una regresión ridge con referencias EMG rezagadas, con y sin su envolvente, se desarrolló con la regularización y la escala de la sustracción como parámetros de su grilla (Sección 3.6) y se evaluó como los demás métodos. Su S_{bal} medio fue menor que el de iCanClean (diferencia de $-3{,}47$ puntos) y no alcanzó la ventaja mínima de un punto del criterio de la Sección 4.5.8. Sus resultados se resumen en el Apéndice F.

En cuanto al tercer objetivo, el AF por bloques, ASR, mediante asrpy, y la CCA con rezagos se implementaron en el mismo entorno. ASR requirió adaptar la biblioteca y retirar un canal dentro del método, por el rango que deja la referencia promedio, y, en E1M1_R2 y E2E3S8, cuánto modificó el reposo dependió de detalles mínimos de la calibración (Sección 5.2.2). Con referencias nulas, la CCA con rezagos devolvió la entrada salvo por el redondeo, mientras que el AF por bloques, por su

término de sesgo, la modificó levemente.

En cuanto al cuarto objetivo, los métodos se evaluaron en reposo y en masticación con la configuración de cada participante elegida con los otros cuatro, mediante métricas, inspección visual, controles, análisis de sensibilidad y una prueba semisintética. ASR obtuvo la mayor reducción media de RMS en masticación, 60,6 %, pero también la mayor modificación del reposo, con un R^2_{reposo} medio de 61,1 %. iCanClean redujo la RMS en masticación en 10,7 % en media, con un R^2_{reposo} medio de 99,4 %, y el AF por bloques y la CCA con rezagos obtuvieron valores del mismo orden, con medias de S_{bal} que no permiten ordenarlos (Sección 5.6.1). En 20–45 Hz, ASR también redujo la potencia, mientras que en los demás métodos la reducción media fue pequeña o negativa. La prueba semisintética y la inspección visual fueron coherentes con estas lecturas.

Respecto del objetivo general, iCanClean, aplicado con referencias EMG de los maseteros, redujo la amplitud del EEG registrado durante la masticación en una magnitud del mismo orden que el AF por bloques y la CCA con rezagos, y modificó poco el reposo, en buena parte porque rara vez retiró componentes en esa condición. Esa reducción se concentró en 1–8 Hz, cuyo origen muscular no está establecido, y no en 20–45 Hz, cuya reducción de potencia es el indicador utilizado más cercano a la actividad muscular. Por ello, la evaluación no permite establecer qué parte de lo retirado correspondía al artefacto.

6.2. Evaluación de la hipótesis y alcance

La hipótesis plantea que iCanClean atenúa más los artefactos de masticación que ASR, AF y CCA, preservando en mayor medida la dinámica espectral y la morfología de la señal en reposo. Con los criterios descriptivos de la Sección 4.5.8, la etiqueta de la hipótesis fue «no respaldada»: el criterio de atenuación en masticación no se cumplió frente a ningún comparador, y el de modificación del reposo se cumplió solo frente a ASR, un veredicto frágil ante las repeticiones de ASR, con las que, sin embargo, la etiqueta no cambió.

Esta etiqueta se lee junto con los controles. La atenuación de iCanClean dependió de sus propias referencias, como la de la CCA con rezagos, y lo que retiró en masticación compartió con el EOG vertical menos varianza en 1–8 Hz que la entrada, como en el AF por bloques y en la CCA, de modo que estos controles no la distinguen de la de esos comparadores. Con la derivación horizontal, en cambio, lo retirado por iCanClean compartió más varianza que la entrada. Además, su menor modificación del reposo frente a ASR refleja, en varios participantes, que no retiró ninguna componente en esa condición (Sección 5.6.4). Por lo tanto, la evaluación no respalda que iCanClean, con estas referencias, atenúe más los artefactos de masticación que los demás métodos. La etiqueta resume esos criterios y no califica la calidad de la limpieza de ninguno de los métodos.

Estas conclusiones se refieren a cinco participantes de E1 y E2, registrados en un mismo laboratorio, a la masticación de los intervalos anotados, a referencias EMG de los maseteros y a las configuraciones elegidas con el procedimiento de esta tesis. Las limitaciones que las acotan, incluidos los registros de masticación no evaluados, se presentan en la Sección 5.8.5.

6.3. Recomendaciones y trabajo futuro

Para evaluar métodos de limpieza del EEG durante la masticación se recomienda:

- informar la atenuación por banda, en particular en 20–45 Hz, porque la reducción de RMS puede quedar dominada por componentes lentas de origen no necesariamente muscular;
- leer la similitud en reposo junto con la razón de RMS y con la actividad del método en esa condición, porque una similitud máxima puede indicar solo que el método no actuó;
- aplicar a los métodos con referencias un control con las de otro participante y otro con referencias nulas, y, en los canales frontales, un control con el EOG;
- en iCanClean con pocas referencias, fijar antes de la evaluación la duración de la ventana y el umbral con el nivel de azar de ρ_1^2 estimado en datos separados, algo que esta tesis no probó;
- en ASR, calibrar con al menos el minuto de reposo recomendado [49] y examinar la sensibilidad a los detalles de la calibración;
- en nuevos registros, registrar el EMG de otros músculos masticatorios, como el temporal, considerar sensores inerciales mandibulares [62] y documentar los bloques de masticación y los códigos de evento.

Como líneas de trabajo futuro se proponen:

- incorporar los bloques de masticación de los archivos `ch` de E1 y los intervalos candidatos de E1M3_R1 (Apéndice G), con sus particiones y reservas fijadas de antemano, y ampliar el número de participantes para permitir un contraste confirmatorio;
- evaluar los métodos con una referencia de la actividad cerebral, como fuentes simuladas, un fantoma [12] o respuestas cerebrales conocidas;
- en iCanClean, examinar ventanas solapadas con una transición gradual, la equivalencia con la implementación oficial y referencias monopolares rerreferenciadas a su propio promedio;
- evaluar combinaciones de métodos, como iCanClean con ASR, junto con su efecto en el reposo y con controles de especificidad.

Apéndices

A. Detalles de implementación y mediciones en los datos de desarrollo

Las mediciones de este apéndice se hicieron con los tramos reservados, que forman el conjunto de desarrollo (Sección 4.3), y no utilizan las ventanas de evaluación.

A.1. Preprocesamiento

Después del filtrado, el cociente entre el valor absoluto de la media y la desviación estándar de cada canal no superó 0,047, y la potencia en 49–51 Hz, en relación con la de 1–45 Hz, no superó $3{,}1 \cdot 10^{-7}$; en la evaluación, un canal primario que supere 0,001 se informa junto con sus métricas. En los registros almacenados, el valor absoluto de la media de los 64 canales EEG en cada muestra se situó entre $1{,}0 \cdot 10^{3}$ y $8{,}3 \cdot 10^{3}$ μV, mientras que con una referencia promedio sería nulo. La regla de los canales de gran amplitud marcó solo el canal FT7 de E1M2_R2 durante la masticación, con una RMS de unas 57 veces la mediana. Los canales EMG monopolares se correlacionaron con el EEG por encima del nivel de azar también durante el reposo, a diferencia de las derivaciones bipolares; es decir, comparten con el EEG actividad que no proviene de la masticación.

En la envolvente, cada rechazo de banda es un filtro recursivo de segundo orden con factor de calidad 30, y los rechazos y el filtro pasa-bajos se aplican hacia adelante y hacia atrás. La media se resta antes del último filtro, que no elimina por completo la componente continua de la envolvente rectificada, y los valores levemente negativos que dejan después de algunas ráfagas los lóbulos del pasa-bajos no se recortan, para no modificar el filtro.

A.2. ASR

La escala global de ASR es

$$s_{\mathrm{ASR}} = \sqrt{\frac{1}{T_{\mathrm{cal}}\,(C-1)} \sum_{n=1}^{T_{\mathrm{cal}}} \sum_{c \neq \mathrm{O2}} \left(x_{\mathrm{cal},c}[n] - \bar{x}_{\mathrm{cal},c}\right)^2}, \tag{A.1}$$

donde T_{cal} es el número de muestras de calibración, la segunda suma recorre los $C - 1 = 63$ canales que recibe ASR y $\bar{x}_{\mathrm{cal},c}$ es la media del canal c en la calibración. Como ASR compara la energía de distintas direcciones espaciales, una sola escala conserva las proporciones entre canales y la geometría de la covarianza, y lleva la señal a un orden de magnitud cercano a la unidad, de modo que las tolerancias numéricas absolutas de la biblioteca no dependen de la unidad de las señales.

De las dos adaptaciones de asrpy, la primera adapta a NumPy 2 la función que ajusta la distribución de amplitudes, sin cambiar su algoritmo. La segunda calcula la pseudoinversa de la

matriz enmascarada $\boldsymbol{D}_w\boldsymbol{\Phi}_w^{\mathsf{T}}\boldsymbol{M}$ de la ecuación (3.15) solo con sus filas activas, para que los residuos numéricos de las filas nulas no se interpreten como direcciones válidas; en aritmética exacta, el resultado es el mismo. Como asrpy y `clean_rawdata` calculan $m = \mathrm{redondeo}((C-1)\,f_{\mathrm{ASR}})$ y la regla de la ecuación (3.14) usa una desigualdad estricta [67, 68], en cada actualización pueden reconstruirse hasta $m+1$ direcciones: 43 con $f_{\mathrm{ASR}} = 0{,}66$ y 22 con $f_{\mathrm{ASR}} = 0{,}33$. Las piezas largas se procesan por bloques con una versión de la función de procesamiento que conserva el estado entre bloques y reproduce la salida de una llamada única en piezas de hasta 30 s.

Retirar otro canal en lugar de O2 cambió de forma apreciable la salida de masticación. Con ASR aplicado a los 64 canales, cambios mínimos de la calibración, como desplazar en una muestra las ventanas con que se estiman los umbrales, modificaron de forma apreciable los umbrales y la salida de masticación, y el intervalo de actualización también cambió la salida cuando ASR actuaba. Como esas sensibilidades se midieron con los 64 canales y no con los 63 que recibe ASR, con la configuración elegida ASR se repite con las variaciones de la Sección 4.5.10.

A.3. CCA con rezagos

Con rezagos de 16 y 32 muestras, equivalentes a 4 y 8 muestras a 256 Hz, el método casi no modificó los datos que no se utilizaron en su ajuste: sus pesos se concentraron en combinaciones de rezagos de varianza casi nula, que solo responden al relleno con ceros de los bordes de cada pieza, y la modificación quedó confinada a esos bordes, dentro de las guardas. Con rezagos de 0, 1 y 2 muestras sí modificó esos datos; por ello, la grilla utiliza estos rezagos, y el de 16 muestras solo aparece en la configuración HIS-CCA (Apéndice B). Fuera de los datos de ajuste, $\alpha = 0{,}01$ resultó pequeño frente a la matriz $\boldsymbol{V}_{:,\mathcal{B}}^{\mathsf{T}}\boldsymbol{V}_{:,\mathcal{B}}$ de la ecuación (3.29) con $L = 0$ y $L = 1$, y también con $L = 2$ en masticación; con $L = 2$ en reposo, la regularización puede reducir parte de la estimación. Con las referencias principales de otro participante, el nivel de azar de $|\rho_i|$ fue cercano a 0,10 en segmentos de 6 a 10 s, de modo que, en las piezas cortas, los umbrales de la grilla se sitúan cerca de ese nivel.

A.4. iCanClean

La implementación oficial selecciona los pares con una desigualdad estricta [60], que difiere de la regla $\rho_i^2 \geq \theta$ solo cuando una correlación coincide exactamente con el umbral; además, incluye ambos extremos de cada ventana y desplaza la última hacia atrás para que quepa en el registro. Estas diferencias de ventaneo cambiaron la salida en a lo sumo un 0,63 % por canal y no cambiaron las componentes seleccionadas. Cuando la última ventana de 4 s se extiende hacia la guarda final, su CCA y su proyección se estiman en parte con esa guarda, cuyo último segundo depende de la extensión por reflexión del filtro. Con referencias de otro participante, la mediana de la mayor correlación canónica al cuadrado, ρ_1^2, fue de 0,64–0,69 con ventanas de 2 s y de 0,45–0,48 con ventanas de 4 s. Así, con ventanas de 2 s, los umbrales de 0,20 a 0,50 quedan bajo la mediana del azar, y con ventanas de 4 s, bajo su percentil 95; solo $\theta = 0{,}65$ lo supera.

A.5. Término de sesgo del AF por bloques

La inclusión del término de sesgo en el AF se decide con una regla aplicada a los datos de desarrollo, que compara dos atenuaciones de potencia en 1–8 Hz, calculadas como $1 - P_{\mathrm{out}}/P_{\mathrm{in}}$ con la ecuación (4.6) y promediadas sobre los diez canales primarios y sobre los cinco participantes: la que produce el AF con sesgo cuando las referencias son nulas y la que produce con las referencias reales. Ambas se miden en la reserva de masticación, con la calibración propia de cada participante y con cada una de las doce configuraciones de la grilla del AF; como la reserva es también el segmento de calibración, la segunda se mide sobre las mismas muestras del ajuste. Una configuración cumple la regla si la primera atenuación alcanza al menos la mitad de la segunda, y si la cumple la mayoría estricta de las configuraciones, el AF no incluye el sesgo. La regla se cumplió en cuatro de las doce configuraciones, las de $\mu = 0{,}1$, de modo que el sesgo se mantiene en toda la grilla. Esta regla mide si el sesgo podría producir por sí solo una atenuación comparable a la del filtro, y no cuánto aporta cuando hay referencias.

Con los valores fijos de la fuga y del avance, la fuga por sí sola tarda unos 5 s en reducir a la mitad los pesos de la calibración, que son el punto de partida de la adaptación; por eso, una calibración con otros participantes condicionaría una parte apreciable de cada pieza de evaluación.

A.6. Controles

En el control (a), los participantes se recorren en el orden cíclico E1M1_R2, E1M2_R2, E2E3S7, E2E3S8, E2E3S9, que después vuelve a E1M1_R2. Las referencias provienen del primer participante siguiente que tenga una pieza de evaluación de la misma condición con longitud suficiente, la primera en el tiempo si hay varias, y se recortan desde el inicio de su pieza a la longitud de la pieza que se limpia. Si ningún participante la tiene, se usa la pieza más larga del participante siguiente, y la pieza que se limpia se recorta desde su inicio a esa longitud, con guardas de 2 s, para compararla con las referencias propias en las mismas ventanas; en la evaluación, ese caso ocurre solo en la pieza de masticación de 58 s de E1M2_R2, cuyo control cubre 23 de sus 27 ventanas. En los datos de desarrollo, el control (a) con la calibración propia aumentó la potencia de 20–45 Hz de la salida de la variante extendida en todos los casos examinados.

En el control (b), los ceros se introducen después de la estandarización, de modo que todos los regresores, incluida la envolvente, valen cero; en la CCA con rezagos, que estandariza las referencias dentro de su propio procedimiento, se introducen antes, en la calibración y en la pieza, y con $\alpha > 0$ la proyección es nula. En el control con el EOG, los canales EOG se centran y se filtran en 1–45 Hz en cada pieza, sin rerreferenciarlos; la entrada, la salida y el EOG se filtran además con un FIR de 1–8 Hz, y la fracción de la varianza que explica el EOG se obtiene con una regresión por mínimos cuadrados con intercepto y sin rezagos.

B. Configuraciones seleccionadas por participante

La Tabla B.1 indica, para cada método y cada participante evaluado, la configuración elegida con los otros cuatro participantes según el procedimiento de la Sección 4.5.7 y el puntaje balanceado S_{bal} que obtiene en el participante evaluado, calculado sobre sus ventanas de evaluación.

Método		E1M1_R2	E1M2_R2	E2E3S7	E2E3S8	E2E3S9
ASR	Configuración	ASR-09	ASR-11	ASR-11	ASR-09	ASR-11
	S_{bal}	79,0	54,5	54,8	60,2	55,8
AF por bloques	Configuración	AF-10	AF-10	AF-10	AF-10	AF-10
	S_{bal}	59,0	58,3	50,4	57,1	54,8
CCA con rezagos	Configuración	CCA-17	CCA-17	CCA-16	CCA-17	CCA-17
	S_{bal}	59,7	50,4	49,9	58,7	55,0
iCanClean	Configuración	ICC-15	ICC-15	ICC-15	ICC-15	ICC-15
	S_{bal}	59,1	52,8	49,3	58,7	55,4
Variante extendida	Configuración	VAR-19	VAR-19	VAR-20	VAR-20	VAR-31
	S_{bal}	56,0	49,4	49,0	57,8	45,8

Tabla B.1: Configuración elegida para cada participante evaluado con los otros cuatro participantes y su S_{bal} en el participante evaluado. Los identificadores remiten a la Tabla B.3.

La Tabla B.2 indica cómo se resolvió cada elección. En ella, la desviación en reposo de una configuración es su media de $|1 - Q_{\mathrm{RMS,reposo}}|$ en los participantes de selección, que decide entre las configuraciones empatadas. En ninguna elección hubo configuraciones no elegibles ni participantes excluidos del promedio. En ASR, todas las elecciones resultaron de un empate entre configuraciones con $k = 20$ o mayor, que la desviación en reposo redujo a un par con la misma tolerancia y distinta fracción f_{ASR}, con las mismas métricas en todos los participantes, de modo que el orden de los identificadores solo eligió entre configuraciones equivalentes. En el AF por bloques y en iCanClean, la mayoría de las elecciones no dependió de un empate, mientras que en la CCA con rezagos el orden de los identificadores eligió, en varias elecciones, entre configuraciones con la misma desviación y distinto puntaje medio. En E2E3S7, la elección de la CCA figura como resuelta por la desviación solo por diferencias de redondeo numérico, que no cambian la configuración elegida.

Tabla B.2: Resolución de la selección de cada método para cada participante evaluado (Sección 4.5.7). Configuraciones empatadas: las que quedan a menos de 0,5 puntos del máximo puntaje medio con los otros cuatro participantes, incluida la elegida; entre paréntesis, su número. Menor desv.: empatadas con la menor desviación en reposo, $|1 - Q_{\mathrm{RMS,reposo}}|$ medio. Mayor: «sí» si la elegida es también la de mayor puntaje medio. Menor desv., Resolución y Mayor se derivan de la selección de la evaluación (Apéndice E).

Participante	Configuraciones empatadas	Menor desv.	Resolución	Mayor
ASR				
E1M1_R2	ASR-07, ASR-08, ASR-09 y ASR-10 (4)	2	por identificador	no
E1M2_R2	ASR-07, ASR-08, ASR-09, ASR-10, ASR-11 y ASR-12 (6)	2	por identificador	no
E2E3S7	ASR-07, ASR-08, ASR-09, ASR-10, ASR-11 y ASR-12 (6)	2	por identificador	sí
E2E3S8	ASR-07, ASR-08, ASR-09 y ASR-10 (4)	2	por identificador	no
E2E3S9	ASR-07, ASR-08, ASR-09, ASR-10, ASR-11 y ASR-12 (6)	2	por identificador	sí
AF por bloques				
E1M1_R2	AF-10 (1)	1	sin empate	sí
E1M2_R2	AF-10 (1)	1	sin empate	sí
E2E3S7	AF-10 y AF-11 (2)	1	por desviación en reposo	no
E2E3S8	AF-10 (1)	1	sin empate	sí
E2E3S9	AF-10 y AF-11 (2)	1	por desviación en reposo	sí
CCA con rezagos				
E1M1_R2	CCA-15, CCA-16, CCA-17, CCA-18, CCA-24, CCA-25, CCA-26 y CCA-27 (8)	2	por identificador	no
E1M2_R2	CCA-06, CCA-08, CCA-09, CCA-15, CCA-16, CCA-17, CCA-18, CCA-24, CCA-25, CCA-26 y CCA-27 (11)	4	por identificador	no
E2E3S7	CCA-13, CCA-16, CCA-22, CCA-25 y CCA-26 (5)	1	por desviación en reposo	no
E2E3S8	CCA-07, CCA-15, CCA-16, CCA-17, CCA-18, CCA-24, CCA-25, CCA-26 y CCA-27 (9)	2	por identificador	no
E2E3S9	CCA-06, CCA-07, CCA-08, CCA-09, CCA-15, CCA-16, CCA-17, CCA-18, CCA-24, CCA-25, CCA-26 y CCA-27 (12)	2	por identificador	no
iCanClean				
E1M1_R2	ICC-15 (1)	1	sin empate	sí
E1M2_R2	ICC-15 (1)	1	sin empate	sí
E2E3S7	ICC-07, ICC-13 y ICC-15 (3)	1	por desviación en reposo	sí
E2E3S8	ICC-15 (1)	1	sin empate	sí
E2E3S9	ICC-13 y ICC-15 (2)	1	por desviación en reposo	sí
Variante extendida				
E1M1_R2	VAR-19, VAR-20, VAR-25, VAR-26, VAR-31 y VAR-32 (6)	1	por desviación en reposo	no
E1M2_R2	VAR-01, VAR-02, VAR-07, VAR-08, VAR-13, VAR-14, VAR-19, VAR-20, VAR-25, VAR-26, VAR-31 y VAR-32 (12)	1	por desviación en reposo	sí
E2E3S7	VAR-19, VAR-20, VAR-25, VAR-26, VAR-31 y VAR-32 (6)	1	por desviación en reposo	sí
E2E3S8	VAR-19, VAR-20, VAR-25, VAR-26, VAR-31 y VAR-32 (6)	1	por desviación en reposo	no
E2E3S9	VAR-19, VAR-20, VAR-25, VAR-26, VAR-31, VAR-32 y VAR-33 (7)	1	por desviación en reposo	sí

La Tabla B.3 presenta los códigos de las configuraciones de cada método, que se utilizan en la Tabla B.1 y en las figuras del Capítulo 5. Las configuraciones equivalentes se calculan una sola vez y se identifican con el código menor.

ASR

k	$f_{\mathrm{ASR}} = 0{,}33$	$f_{\mathrm{ASR}} = 0{,}66$
5	ASR-01	ASR-02
10	ASR-03	ASR-04
15	ASR-05	ASR-06
20	ASR-07	ASR-08
25	ASR-09	ASR-10
30	ASR-11	ASR-12

AF por bloques

Envolvente	L_{AF} (muestras)	$\mu = 0{,}01$	$\mu = 0{,}03$	$\mu = 0{,}1$
sin	64	AF-01	AF-02	AF-03
sin	256	AF-04	AF-05	AF-06
con	64	AF-07	AF-08	AF-09
con	256	AF-10	AF-11	AF-12

CCA con rezagos

n_{comp}	L (muestras)	$\tau = 0{,}10$	$\tau = 0{,}15$	$\tau = 0{,}20$
2	0	CCA-01	CCA-02	CCA-03
2	1	CCA-04	CCA-05	CCA-06
2	2	CCA-07	CCA-08	CCA-09
4	0	CCA-10 (= CCA-01)	CCA-11 (= CCA-02)	CCA-12 (= CCA-03)
4	1	CCA-13	CCA-14	CCA-15
4	2	CCA-16	CCA-17	CCA-18
6	0	CCA-19 (= CCA-01)	CCA-20 (= CCA-02)	CCA-21 (= CCA-03)
6	1	CCA-22	CCA-23	CCA-24
6	2	CCA-25	CCA-26	CCA-27

iCanClean

Ventana (s)	θ	Fuente $\boldsymbol{U}$	Fuente $\boldsymbol{V}$
2	0,20	ICC-01	ICC-02
2	0,35	ICC-03	ICC-04
2	0,50	ICC-05	ICC-06
2	0,65	ICC-07	ICC-08
4	0,20	ICC-09	ICC-10
4	0,35	ICC-11	ICC-12
4	0,50	ICC-13	ICC-14
4	0,65	ICC-15	ICC-16
8	0,20	ICC-W8-01	ICC-W8-02
8	0,35	ICC-W8-03	ICC-W8-04
8	0,50	ICC-W8-05	ICC-W8-06
8	0,65	ICC-W8-07	ICC-W8-08

Variante extendida

		Sin envolvente		Con envolvente	
λ	γ	$L = 16$	$L = 32$	$L = 16$	$L = 32$
0,001	0,6	VAR-01	VAR-02	VAR-19	VAR-20
0,001	0,8	VAR-03	VAR-04	VAR-21	VAR-22
0,001	1,0	VAR-05	VAR-06	VAR-23	VAR-24
0,01	0,6	VAR-07	VAR-08	VAR-25	VAR-26
0,01	0,8	VAR-09	VAR-10	VAR-27	VAR-28
0,01	1,0	VAR-11	VAR-12	VAR-29	VAR-30
0,1	0,6	VAR-13	VAR-14	VAR-31	VAR-32
0,1	0,8	VAR-15	VAR-16	VAR-33	VAR-34
0,1	1,0	VAR-17	VAR-18	VAR-35	VAR-36

Tabla B.3: Identificadores de las configuraciones de cada método: en ASR, según la tolerancia k y la fracción dimensional f_{ASR}; en el AF por bloques, según la envolvente, la memoria L_{AF} y el paso μ; en la CCA con rezagos, según el número de componentes n_{comp}, el rezago máximo L y el umbral τ; en iCanClean, según la duración de la ventana, el umbral θ y la fuente de la actividad que se resta; y en la variante extendida, según la regularización λ, la escala γ, la envolvente y el rezago máximo L, en muestras. Con $L = 0$, las configuraciones de la CCA con $n_{\mathrm{comp}} = 4$ y $n_{\mathrm{comp}} = 6$ equivalen a las de $n_{\mathrm{comp}} = 2$ y se identifican con el código de estas. Las configuraciones de iCanClean de 8 s no forman parte de la grilla principal y se evalúan solo en la sensibilidad a la duración de la ventana (Sección 4.5.10).

Además de las configuraciones elegidas, se informan sin selección la identidad (Tabla 5.7), configuraciones de referencia y configuraciones adicionales (Tabla B.4; las métricas de HIS-VAR, en la Tabla F.1). Las de referencia son LIT-ICC-4, la óptima que Gonsisko et al. obtuvieron con muchos más canales de referencia [58]; LIT-ICC-2, con los valores predeterminados de la implementación oficial [60]; y LIT-ASR, con los de asrpy [68]. Las adicionales, HIS-ASR, HIS-AF, HIS-CCA e HIS-VAR, pudieron elegirse con datos de E1M1_R2, sin que pueda descartarse que interviniera otro participante; por ello, E1M1_R2 se excluye de su resumen. Todas se aplican con el preprocesamiento, las referencias y los parámetros fijos de esta evaluación, y HIS-CCA, con $L = 16$ muestras, e HIS-VAR, con $\lambda = 0{,}01/T_{\mathrm{cal}}$, quedan fuera de la grilla de la Tabla 4.2.

Rótulo	Método	Configuración y origen	Equivalencia
LIT-ASR	ASR	$k = 20$, $f_{\mathrm{ASR}} = 0{,}66$; valores predeterminados de asrpy [68]	ASR-08
HIS-ASR	ASR	$k = 20$, $f_{\mathrm{ASR}} = 0{,}33$; configuración adicional	ASR-07
HIS-AF	AF por bloques	con envolvente, $L_{\mathrm{AF}} = 256$ muestras, $\mu = 0{,}1$; configuración adicional	AF-12
HIS-CCA	CCA con rezagos	$n_{\mathrm{comp}} = 6$, $L = 16$ muestras, $\tau = 0{,}10$; configuración adicional	fuera de la grilla
LIT-ICC-4	iCanClean	ventana de 4 s, $\theta = 0{,}65$, fuente $\boldsymbol{U}$; Gonsisko et al. [58]	ICC-15
LIT-ICC-2	iCanClean	ventana de 2 s, $\theta = 0{,}85$, fuente $\boldsymbol{U}$; implementación oficial [60]	fuera de la grilla
HIS-VAR	Variante extendida	sin envolvente, $\lambda = 0{,}01/T_{\mathrm{cal}}$, $\gamma = 0{,}8$, $L = 32$ muestras; configuración adicional	fuera de la grilla

	Reposo		Masticación		Puntaje	
Configuración	R^2	Q_{RMS}	ΔRMS	Reducción en 20–45 Hz	S_{bal}	$S_{\mathrm{bal,banda}}$
LIT-ASR	58,0 %	0,727	64,4 %	68,3 %	61,2	63,2
HIS-ASR	51,1 %	0,675	63,3 %	68,8 %	57,2	60,0
HIS-AF	74,7 %	0,893	24,4 %	3,35 %	49,5	39,0
HIS-CCA	100,0 %	1,00	0,04 %	0,02 %	50,0	50,0
LIT-ICC-4	99,4 %	0,997	10,7 %	−3,51 %	55,1	48,0
LIT-ICC-2	99,5 %	0,997	9,03 %	−4,24 %	54,3	47,6

Tabla B.4: Configuraciones informadas sin selección. Arriba, su equivalencia con los identificadores de la Tabla B.3; abajo, la media de los participantes de las métricas principales y de los puntajes, con las columnas de la Tabla 5.7, que presenta los métodos con la configuración elegida. Las configuraciones HIS se resumen sin E1M1_R2.

LIT-ASR e HIS-ASR, que difieren solo en f_{ASR}, obtuvieron los mismos valores en cada participante, y sus medias difieren solo porque la de HIS-ASR se calcula sin E1M1_R2. HIS-CCA prácticamente no modificó la señal. HIS-AF, que difiere de la elegida del AF por bloques solo en el paso de adaptación, el mayor de la grilla, redujo más la RMS en masticación, pero también modificó más el reposo: en E1M2_R2, por ejemplo, su reducción de RMS y su R^2_{reposo} fueron 34,6 % y 73,7 %, frente a 18,6 % y 98,0 % con la configuración elegida.

C. Resultados complementarios

C.1. Verificaciones del preprocesamiento

Estas verificaciones (Sección 4.2) no modifican el preprocesamiento. La continua residual, es decir, el valor absoluto de la media de la zona de evaluación dividido por su desviación estándar, se calculó en 832 pares de canal y pieza, los 64 canales de las 13 piezas de evaluación; su mediana fue 0,00213, su percentil 95, 0,0247, y su máximo, 0,136. El umbral de 0,05 se superó en 5 pares, todos en la pieza de masticación de E1M1_R2 comprendida entre la reserva de calibración y el corte, que contiene una sola ventana de evaluación, la que muestran las trazas de masticación de la inspección visual. Los canales que lo superaron fueron F1, C5, CP3, AF4 y C2, y de ellos solo AF4 es un canal primario. En esa pieza, la mediana de la continua residual de los 64 canales fue 0,0260, frente a 0,0106 en la otra pieza de masticación del mismo participante. Como las métricas centran cada ventana y la PSD elimina la media de cada segmento, ese residuo no interviene directamente en ellas, aunque forma parte de la entrada de los métodos.

La potencia en 49–51 Hz después del filtro, relativa a la de 1–45 Hz, tuvo una mediana de $3{,}07 \cdot 10^{-8}$ y un máximo de $1{,}81 \cdot 10^{-7}$ en los 64 canales, y un máximo de $1{,}63 \cdot 10^{-7}$ en los canales primarios, muy por debajo del umbral de 0,001.

La regla de los canales de gran amplitud marcó 4 pares de canal y pieza en 3 de las 13 piezas de evaluación, todos en las piezas de masticación de E1M2_R2: FT7 y PO3 en la primera, FT7 en la segunda y FT7 en la tercera. La RMS de FT7 fue 63,4, 42,7 y 37,5 veces la mediana, respectivamente, frente a un umbral de 20 veces. En las demás piezas, la mayor RMS relativa fue la de FC5 en la masticación de E2E3S9, 19,4 veces la mediana, cercana al umbral, pero sin superarlo. Los canales marcados se conservaron en el promedio de referencia, y su efecto se examina en el Apéndice D.3.

C.2. Discontinuidades de iCanClean entre ventanas

La Tabla C.1 resume las discontinuidades de iCanClean en los bordes entre ventanas consecutivas (Sección 4.4.4). En E1M1_R2 y E2E3S7, la mediana del salto fue 0 y 0, porque la mayoría de los bordes separan ventanas sin componentes retiradas. LIT-ICC-2, con ventanas de 2 s, tiene alrededor del doble de bordes, y su salto máximo fue mayor que el de la configuración elegida en E2E3S8 y E2E3S9.

C.3. Potencia por banda por participante

Por participante (Tabla C.2), E2E3S7 fue en masticación la excepción de ASR, coherente con su menor reducción de RMS en ese participante, y la CCA con rezagos e iCanClean aumentaron la potencia de beta y gamma en E1M2_R2. En reposo, iCanClean no modificó ninguna banda en los participantes en que no retiró componentes.

	Configuración elegida			LIT-ICC-2		
Participante	Bordes	Mediana	Máximo	Bordes	Mediana	Máximo
E1M1_R2	3	0	0	8	0	0
E1M2_R2	30	0,109	1,21	61	0,107	1,18
E2E3S7	16	0	0,575	32	0	0,277
E2E3S8	16	0,105	0,660	32	0,108	1,30
E2E3S9	16	0,0936	0,861	32	0,0320	2,08

Tabla C.1: iCanClean: discontinuidades en los bordes entre ventanas consecutivas de una misma pieza, en ambas condiciones, con la configuración elegida para cada participante, igual a LIT-ICC-4, y con LIT-ICC-2: número de bordes y mediana y máximo del salto de la estimación del artefacto, relativo a la RMS de la entrada en el par de ventanas. La mediana se calcula sobre los bordes, con la mediana de los canales primarios en cada uno, y el máximo, sobre todos los bordes y canales primarios.

Tabla C.2: Razón de potencia $P_{\mathrm{out}}/P_{\mathrm{in}}$ por banda (ecuación (4.6)) en reposo y en masticación, para cada método con la configuración elegida para cada participante: valor de cada participante y media. Para la identidad, que devuelve la entrada, se indica solo la media. Una razón menor que uno indica que la potencia de la banda se atenuó.

Método	Participante	Delta 1–4 Hz	Theta 4–8 Hz	Alfa 8–13 Hz	Beta 13–30 Hz	Gamma 30–45 Hz
Reposo						
Identidad	Media	1,00	1,00	1,00	1,00	1,00
ASR	E1M1_R2	0,917	1,01	0,749	0,931	0,995
	E1M2_R2	0,116	0,193	0,591	0,846	0,947
	E2E3S7	1,00	1,00	1,00	1,00	1,00
	E2E3S8	0,331	0,451	0,503	0,971	0,985
	E2E3S9	0,404	0,434	0,749	0,976	0,974
	Media	0,554	0,618	0,719	0,945	0,980
AF por bloques	E1M1_R2	0,952	1,00	0,991	1,00	1,00
	E1M2_R2	0,920	0,973	0,999	1,01	1,01
	E2E3S7	0,861	0,953	1,02	1,01	1,00
	E2E3S8	0,956	0,990	0,995	1,00	1,00
	E2E3S9	0,949	1,01	1,03	1,00	1,01
	Media	0,928	0,986	1,01	1,00	1,00
CCA con rezagos	E1M1_R2	0,999	1,00	0,996	0,996	1,01
	E1M2_R2	0,964	0,990	1,01	1,04	1,16
	E2E3S7	0,925	0,941	1,05	1,03	1,03
	E2E3S8	0,999	0,997	0,999	0,999	1,00
	E2E3S9	0,996	0,995	1,01	1,00	0,997
	Media	0,977	0,985	1,01	1,01	1,04
iCanClean	E1M1_R2	1,00	1,00	1,00	1,00	1,00
	E1M2_R2	1,00	1,00	1,00	1,00	1,00
	E2E3S7	0,960	0,990	1,05	1,01	1,02
	E2E3S8	1,00	1,00	1,00	1,00	1,00
	E2E3S9	0,998	0,999	1,00	0,998	0,999
	Media	0,992	0,998	1,01	1,00	1,00

Continúa en la página siguiente

Tabla C.2, continuación

Método	Participante	Delta 1–4 Hz	Theta 4–8 Hz	Alfa 8–13 Hz	Beta 13–30 Hz	Gamma 30–45 Hz
Masticación						
Identidad	Media	1,00	1,00	1,00	1,00	1,00
ASR	E1M1_R2	0,0496	0,222	0,680	0,497	0,296
	E1M2_R2	0,0033	0,0433	0,147	0,157	0,122
	E2E3S7	0,771	0,861	0,906	0,900	0,899
	E2E3S8	0,0349	0,146	0,201	0,153	0,140
	E2E3S9	0,0294	0,163	0,304	0,274	0,147
	Media	0,178	0,287	0,448	0,396	0,321
AF por bloques	E1M1_R2	0,589	1,01	1,03	0,997	0,975
	E1M2_R2	0,559	1,00	1,02	0,997	0,991
	E2E3S7	0,758	0,966	0,990	0,996	0,994
	E2E3S8	0,459	0,936	0,982	0,991	0,994
	E2E3S9	0,631	0,958	0,989	0,992	0,976
	Media	0,599	0,974	1,00	0,995	0,986
CCA con rezagos	E1M1_R2	0,620	0,882	1,08	1,03	0,922
	E1M2_R2	0,929	1,10	1,07	1,06	1,12
	E2E3S7	0,949	0,908	0,923	0,950	0,947
	E2E3S8	0,528	0,725	0,743	0,926	0,933
	E2E3S9	0,824	0,879	1,00	0,947	0,815
	Media	0,770	0,899	0,964	0,982	0,947
iCanClean	E1M1_R2	0,658	0,917	1,00	1,05	0,922
	E1M2_R2	0,828	1,08	1,13	1,25	1,49
	E2E3S7	0,994	0,976	0,954	0,965	0,972
	E2E3S8	0,584	0,698	0,736	0,963	0,873
	E2E3S9	0,814	0,854	0,952	0,900	0,887
	Media	0,776	0,905	0,954	1,03	1,03

C.4. Controles de especificidad por participante

La Tabla C.3 se lee en la Sección 5.6.3.

C.5. Evaluación semisintética por participante

Por participante (Tabla C.4), la CCA con rezagos obtuvo un error menor y un CC mayor que la identidad en todos los participantes y niveles; en el AF por bloques, la excepción fue E2E3S7 con $\mathrm{SNR} = 5$ dB, y en iCanClean, E2E3S9, con un CC menor que el de la identidad en los tres niveles y un $\mathrm{RRMSE_t}$ levemente mayor con $\mathrm{SNR} = 5$ dB. El CC de ASR fue menor que el de la identidad en E1M2_R2, E2E3S8 y E2E3S9 en todos los niveles.

		$\Delta\mathrm{RMS}_{\mathrm{mast}}$				Reducción en 20–45 Hz			
Método	Participante	Propias	(a)	(a′)	(b)	Propias	(a)	(a′)	(b)
AF por bloques	E1M1_R2	19,0 %	4,14 %	3,97 %	0,34 %	1,77 %	−1,07 %	−0,33 %	−0,80 %
	E1M2_R2	18,6 %	5,76 %	6,09 %	0,26 %	1,01 %	−0,40 %	−0,44 %	−0,81 %
	E2E3S7	4,78 %	1,21 %	1,41 %	−0,16 %	0,49 %	−0,44 %	−0,47 %	−0,81 %
	E2E3S8	14,7 %	5,13 %	4,83 %	−0,03 %	0,50 %	−0,58 %	−0,37 %	−0,81 %
	E2E3S9	13,3 %	3,83 %	3,79 %	0,05 %	1,89 %	−0,48 %	−0,39 %	−0,80 %
	Media	14,1 %	4,01 %	4,02 %	0,09 %	1,13 %	−0,59 %	−0,40 %	−0,81 %
CCA con rezagos	E1M1_R2	19,6 %	0,05 %	0,01 %	0 %	6,82 %	−0,87 %	−0,66 %	0 %
	E1M2_R2	2,90 %	0,12 %	0,11 %	0 %	−8,73 %	−0,65 %	−0,08 %	0 %
	E2E3S7	2,94 %	0,11 %	0,22 %	0 %	4,87 %	−0,13 %	−0,13 %	0 %
	E2E3S8	17,4 %	0,12 %	0,12 %	0 %	3,29 %	−0,09 %	−0,02 %	0 %
	E2E3S9	10,1 %	0,15 %	0,17 %	0 %	14,7 %	−0,88 %	−0,18 %	0 %
	Media	10,6 %	0,11 %	0,13 %	0 %	4,18 %	−0,52 %	−0,21 %	0 %
iCanClean	E1M1_R2	18,3 %	0 %	—	0 %	0,79 %	0 %	—	0 %
	E1M2_R2	5,57 %	0 %	—	0 %	−40,8 %	0 %	—	0 %
	E2E3S7	1,38 %	0 %	—	0 %	2,83 %	0 %	—	0 %
	E2E3S8	17,4 %	0,18 %	—	0 %	7,40 %	−0,08 %	—	0 %
	E2E3S9	10,9 %	1,39 %	—	0 %	12,2 %	−1,28 %	—	0 %
	Media	10,7 %	0,31 %	—	0 %	−3,51 %	−0,27 %	—	0 %

		R^2_{reposo}				$Q_{\mathrm{RMS,reposo}}$			
Método	Participante	Propias	(a)	(a′)	(b)	Propias	(a)	(a′)	(b)
AF por bloques	E1M1_R2	99,0 %	98,8 %	98,8 %	99,5 %	0,991	0,991	0,990	1,00
	E1M2_R2	98,0 %	97,5 %	97,1 %	99,1 %	0,977	0,986	0,988	1,00
	E2E3S7	95,9 %	98,5 %	98,8 %	98,8 %	0,960	0,980	0,985	0,998
	E2E3S8	99,4 %	99,1 %	98,3 %	99,6 %	0,991	0,992	0,994	1,00
	E2E3S9	96,2 %	96,7 %	98,1 %	99,0 %	0,985	0,987	0,979	0,999
	Media	97,7 %	98,1 %	98,2 %	99,2 %	0,981	0,987	0,987	1,00
CCA con rezagos	E1M1_R2	99,8 %	99,9 %	99,8 %	100,0 %	0,998	1,00	1,00	1,00
	E1M2_R2	97,9 %	99,8 %	99,9 %	100,0 %	0,990	0,999	0,999	1,00
	E2E3S7	96,8 %	100,0 %	100,0 %	100,0 %	0,972	1,00	1,00	1,00
	E2E3S8	100,0 %	100,0 %	100,0 %	100,0 %	0,999	1,00	1,00	1,00
	E2E3S9	99,9 %	99,5 %	99,1 %	100,0 %	0,999	0,996	0,995	1,00
	Media	98,9 %	99,8 %	99,8 %	100,0 %	0,992	0,999	0,999	1,00
iCanClean	E1M1_R2	100,0 %	100,0 %	—	100,0 %	1,00	1,00	—	1,00
	E1M2_R2	100,0 %	100,0 %	—	100,0 %	1,00	1,00	—	1,00
	E2E3S7	97,2 %	100,0 %	—	100,0 %	0,986	1,00	—	1,00
	E2E3S8	100,0 %	99,9 %	—	100,0 %	1,00	0,999	—	1,00
	E2E3S9	99,9 %	100,0 %	—	100,0 %	0,999	1,00	—	1,00
	Media	99,4 %	100,0 %	—	100,0 %	0,997	1,00	—	1,00

Tabla C.3: Controles de especificidad (Sección 4.5.7), con la configuración elegida para cada participante: valor de cada participante y media. Arriba, masticación; abajo, reposo. Las referencias propias se evalúan en todas las ventanas, y los controles (a) y (a′), en las que cubre el control (a) (Tabla 5.8). En iCanClean, (a′) no se aplica (—).

		SNR $= -5$ dB			SNR $= 0$ dB			SNR $= 5$ dB		
Método	Participante	$\mathrm{RRMSE_t}$	$\mathrm{RRMSE_f}$	CC	$\mathrm{RRMSE_t}$	$\mathrm{RRMSE_f}$	CC	$\mathrm{RRMSE_t}$	$\mathrm{RRMSE_f}$	CC
Identidad	E1M1_R2	1,65	6,11	0,533	0,928	1,93	0,736	0,522	0,619	0,884
	E1M2_R2	2,46	9,97	0,453	1,38	3,17	0,644	0,777	1,01	0,802
	E2E3S7	2,05	3,99	0,444	1,15	1,19	0,660	0,647	0,350	0,831
	E2E3S8	1,45	4,32	0,649	0,815	1,38	0,806	0,458	0,450	0,905
	E2E3S9	1,75	7,07	0,639	0,981	2,27	0,775	0,552	0,745	0,873
	Media	1,87	6,29	0,544	1,05	1,99	0,724	0,591	0,636	0,859
ASR	E1M1_R2	0,989	1,06	0,628	0,711	0,532	0,780	0,576	0,617	0,852
	E1M2_R2	1,15	0,951	0,175	1,11	0,938	0,222	1,02	0,860	0,287
	E2E3S7	1,70	1,57	0,426	1,07	0,770	0,665	0,640	0,337	0,834
	E2E3S8	0,967	0,730	0,441	0,910	0,830	0,536	0,817	0,710	0,606
	E2E3S9	1,07	0,814	0,297	0,975	0,741	0,392	0,924	0,695	0,442
	Media	1,18	1,03	0,394	0,956	0,762	0,519	0,795	0,644	0,604
AF por bloques	E1M1_R2	1,35	3,76	0,605	0,775	1,25	0,787	0,452	0,429	0,905
	E1M2_R2	2,08	5,49	0,494	1,20	1,69	0,673	0,705	0,523	0,814
	E2E3S7	1,95	3,05	0,461	1,11	0,926	0,666	0,650	0,295	0,826
	E2E3S8	1,25	2,31	0,687	0,719	0,740	0,828	0,422	0,253	0,914
	E2E3S9	1,51	3,24	0,649	0,875	1,02	0,783	0,530	0,361	0,876
	Media	1,63	3,57	0,579	0,935	1,12	0,747	0,552	0,372	0,867
CCA con rezagos	E1M1_R2	1,32	3,74	0,622	0,749	1,27	0,804	0,424	0,471	0,919
	E1M2_R2	2,32	8,79	0,470	1,31	2,77	0,659	0,750	0,913	0,810
	E2E3S7	1,95	3,83	0,456	1,10	1,14	0,670	0,623	0,324	0,838
	E2E3S8	1,19	2,37	0,700	0,683	0,796	0,839	0,402	0,293	0,920
	E2E3S9	1,60	5,35	0,642	0,904	1,70	0,781	0,514	0,549	0,880
	Media	1,68	4,82	0,578	0,949	1,54	0,751	0,543	0,510	0,873
iCanClean	E1M1_R2	1,37	3,88	0,599	0,777	1,29	0,789	0,445	0,470	0,910
	E1M2_R2	2,34	8,27	0,465	1,32	2,61	0,653	0,759	0,818	0,804
	E2E3S7	1,97	4,00	0,454	1,12	1,11	0,670	0,638	0,341	0,834
	E2E3S8	1,19	3,32	0,693	0,671	1,06	0,838	0,385	0,348	0,922
	E2E3S9	1,62	5,98	0,629	0,932	1,88	0,768	0,557	0,605	0,868
	Media	1,70	5,09	0,568	0,964	1,59	0,744	0,557	0,516	0,868

Tabla C.4: Evaluación semisintética: $\mathrm{RRMSE_t}$, $\mathrm{RRMSE_f}$ y CC frente a $\boldsymbol{X}_{\mathrm{rep}}$ (ecuación (4.12)), para cada método con la configuración elegida para cada participante, sin reajustarla, y para la identidad, en cada nivel de SNR (ecuación (4.11)): valor de cada participante y media.

D. Análisis de sensibilidad

D.1. Repeticiones de ASR

La Tabla D.1 compara el valor principal de cada participante con el intervalo de las 6 repeticiones de ASR de la Sección 4.5.10. En E1M1_R2 y E2E3S8, las repeticiones modificaron de forma apreciable las métricas, sobre todo en reposo; en los demás participantes, poco. En E1M1_R2, R^2_{reposo} fue 76,3 % con la calibración desplazada una muestra y 100,0 % con O1 retirado, con el que ASR redujo la RMS en masticación en 55,3 % y dejó $Q_{\mathrm{RMS,reposo}}$ en 1,00; en E2E3S8, fue 41,2 % con el mayor desplazamiento y 69,1 % con O1 retirado. Así, en ellos, cuánto modifica ASR el reposo depende de detalles de la calibración ajenos a la configuración elegida.

	S_{bal}		$\Delta\mathrm{RMS}_{\mathrm{mast}}$		R^2_{reposo}		$Q_{\mathrm{RMS,reposo}}$		
Participante	Principal	Repeticiones	Principal	Repeticiones	Principal	Repeticiones	Principal	Repeticiones	Fallos
E1M1_R2	79,0	71,2–81,7	65,5 %	55,3 %–66,4 %	92,5 %	76,3 %–100,0 %	0,964	0,904–1,00	0
E1M2_R2	54,5	54,1–54,8	85,8 %	83,9 %–86,2 %	23,2 %	23,2 %–24,5 %	0,452	0,453–0,465	0
E2E3S7	54,8	54,4–56,0	9,58 %	8,82 %–12,1 %	100,0 %	100,0 %–100,0 %	1,00	1,00–1,00	0
E2E3S8	60,2	56,3–67,7	71,3 %	66,4 %–71,3 %	49,0 %	41,2 %–69,1 %	0,730	0,691–0,836	0
E2E3S9	55,8	55,3–56,4	70,8 %	68,1 %–70,8 %	40,7 %	40,9 %–42,6 %	0,587	0,588–0,600	0

Tabla D.1: ASR: valor principal de cada participante, con la configuración elegida con los otros cuatro participantes, e intervalo entre el mínimo y el máximo de las repeticiones de la Sección 4.5.10, calculado sin el valor principal, que puede quedar fuera de él. Fallos: calibraciones fallidas. $\Delta\mathrm{RMS}_{\mathrm{mast}}$ y R^2_{reposo}, en porcentaje.

D.2. Referencia EMG

Con las referencias alternativas de la Sección 4.2 (Tabla D.2), iCanClean no modificó la señal en ningún participante cuando estaban filtradas en 20–120 Hz, bipolares o monopolares; con los monopolares de 1–45 Hz, en cambio, atenuó más que con la referencia principal en todos los participantes, pero también modificó más el reposo, y esa atenuación no puede atribuirse solo al artefacto. El AF por bloques atenuó más y modificó más el reposo con los monopolares, y menos con las referencias de 20–120 Hz, con las que la CCA con rezagos concentró la atenuación en 20–45 Hz.

Tabla D.2: Sensibilidad a la referencia EMG: en masticación, $\Delta\mathrm{RMS}_{\mathrm{mast}}$ y reducción de potencia en 20–45 Hz, y en reposo, R^2_{reposo} y $Q_{\mathrm{RMS,reposo}}$, del AF por bloques, de la CCA con rezagos y de iCanClean con cada referencia alternativa de la Sección 4.2 y la configuración elegida para cada participante; con la referencia principal, en las Tablas 5.3, 5.4 y 5.6. «Promedio» y «CMS» indican la referencia de los canales monopolares de 1–45 Hz. Valor de cada participante; $\Delta\mathrm{RMS}_{\mathrm{mast}}$, la reducción de potencia y R^2_{reposo}, en porcentaje.

Método	Referencia	E1M1_R2	E1M2_R2	E2E3S7	E2E3S8	E2E3S9
$\Delta\mathrm{RMS}_{\mathrm{mast}}$						
AF por bloques	Monopolares, promedio	23,6 %	24,4 %	7,80 %	15,5 %	15,0 %
	Monopolares, CMS	20,1 %	23,8 %	5,96 %	14,8 %	14,5 %
	Bipolares, 20–120 Hz	13,0 %	18,2 %	3,63 %	12,6 %	9,97 %
	Monopolares, 20–120 Hz	14,2 %	17,9 %	3,84 %	11,3 %	11,4 %
CCA con rezagos	Monopolares, promedio	0,46 %	11,5 %	2,36 %	7,13 %	14,9 %
	Monopolares, CMS	0,17 %	11,5 %	1,03 %	9,07 %	14,0 %
	Bipolares, 20–120 Hz	1,07 %	0,69 %	0,91 %	0,77 %	2,41 %
	Monopolares, 20–120 Hz	1,33 %	2,51 %	2,85 %	2,62 %	3,66 %
iCanClean	Monopolares, promedio	35,5 %	25,9 %	14,3 %	26,4 %	21,1 %
	Monopolares, CMS	28,6 %	22,7 %	10,7 %	26,5 %	21,5 %
	Bipolares, 20–120 Hz	0 %	0 %	0 %	0 %	0 %
	Monopolares, 20–120 Hz	0 %	0 %	0 %	0 %	0 %
Reducción de potencia en 20–45 Hz						
AF por bloques	Monopolares, promedio	0,01 %	2,65 %	0,76 %	1,55 %	5,13 %
	Monopolares, CMS	0,16 %	1,60 %	0,79 %	1,30 %	5,18 %
	Bipolares, 20–120 Hz	0,85 %	1,15 %	0,11 %	0,16 %	0,89 %
	Monopolares, 20–120 Hz	0,93 %	3,35 %	1,41 %	1,19 %	3,63 %

Continúa en la página siguiente

Tabla D.2, continuación

Método	Referencia	E1M1_R2	E1M2_R2	E2E3S7	E2E3S8	E2E3S9
Reducción de potencia en 20–45 Hz (continuación)						
CCA con rezagos	Monopolares, promedio	0,37 %	−5,37 %	3,53 %	7,38 %	25,4 %
	Monopolares, CMS	0,02 %	1,74 %	2,16 %	11,0 %	24,3 %
	Bipolares, 20–120 Hz	18,9 %	9,59 %	5,71 %	7,92 %	17,7 %
	Monopolares, 20–120 Hz	21,3 %	30,9 %	16,1 %	24,0 %	29,3 %
iCanClean	Monopolares, promedio	−0,68 %	−63,1 %	13,8 %	27,4 %	22,5 %
	Monopolares, CMS	−2,41 %	−41,5 %	15,8 %	28,4 %	28,0 %
	Bipolares, 20–120 Hz	0 %	0 %	0 %	0 %	0 %
	Monopolares, 20–120 Hz	0 %	0 %	0 %	0 %	0 %
R^2_{reposo}						
AF por bloques	Monopolares, promedio	99,0 %	94,3 %	84,7 %	96,6 %	95,9 %
	Monopolares, CMS	98,2 %	96,5 %	94,2 %	93,7 %	97,3 %
	Bipolares, 20–120 Hz	99,9 %	99,7 %	99,8 %	99,8 %	99,8 %
	Monopolares, 20–120 Hz	99,8 %	99,8 %	99,5 %	100,0 %	99,9 %
CCA con rezagos	Monopolares, promedio	100,0 %	100,0 %	99,9 %	100,0 %	99,1 %
	Monopolares, CMS	100,0 %	100,0 %	100,0 %	100,0 %	99,9 %
	Bipolares, 20–120 Hz	100,0 %	100,0 %	100,0 %	100,0 %	100,0 %
	Monopolares, 20–120 Hz	99,9 %	100,0 %	99,7 %	99,9 %	100,0 %
iCanClean	Monopolares, promedio	91,4 %	41,8 %	42,2 %	75,0 %	45,5 %
	Monopolares, CMS	84,4 %	80,0 %	70,7 %	75,6 %	77,6 %
	Bipolares, 20–120 Hz	100,0 %	100,0 %	100,0 %	100,0 %	100,0 %
	Monopolares, 20–120 Hz	100,0 %	100,0 %	100,0 %	100,0 %	100,0 %
$Q_{\text{RMS,reposo}}$						
AF por bloques	Monopolares, promedio	0,979	0,851	0,783	0,941	0,904
	Monopolares, CMS	0,966	0,918	0,895	0,921	0,939
	Bipolares, 20–120 Hz	1,00	0,997	0,995	0,999	0,999
	Monopolares, 20–120 Hz	0,996	0,998	0,988	1,00	0,997
CCA con rezagos	Monopolares, promedio	1,00	0,992	0,977	1,00	0,926
	Monopolares, CMS	0,994	0,993	1,00	0,997	0,993
	Bipolares, 20–120 Hz	1,00	1,00	1,00	1,00	1,00
	Monopolares, 20–120 Hz	0,998	0,999	0,996	1,00	1,00
iCanClean	Monopolares, promedio	0,955	0,656	0,641	0,860	0,648
	Monopolares, CMS	0,917	0,901	0,830	0,868	0,885
	Bipolares, 20–120 Hz	1,00	1,00	1,00	1,00	1,00
	Monopolares, 20–120 Hz	1,00	1,00	1,00	1,00	1,00

D.3. Decisiones de diseño

La Tabla D.3 presenta el criterio 1 y el requisito previo calculados solo con las 47 ventanas de masticación independientes del diseño (Sección 4.5.10), de 4 participantes. Frente a ASR y a la CCA con rezagos, el criterio 1 dio los mismos resultados que con todas las ventanas, y frente al AF por bloques se cumplió en E2E3S8 y E2E3S9; el requisito previo terminó en empate en E2E3S7.

Sin la primera ventana evaluada de cada pieza con reserva (Tablas D.4 y D.5), la $\Delta\mathrm{RMS}_{\mathrm{mast}}$ de E1M1_R2 disminuyó en todos los métodos, hasta anularse en iCanClean, y el S_{bal} medio bajó en todos ellos. En E1M1_R2, el requisito previo pasó a empate, con una diferencia de 0 puntos porcentuales, y la comparación con el AF en el criterio 1, a no cumplida (−6,64 puntos porcentuales

Comparación	Participante	Ventanas	$\Delta\mathrm{RMS}_{\mathrm{mast}}$	20–45 Hz	Veredicto
Requisito previo	E1M1_R2	0	sin ventanas de este tipo		
	E1M2_R2	17	5,60	—	cumple
	E2E3S7	10	0	—	empate
	E2E3S8	10	16,6	—	cumple
	E2E3S9	10	18,1	—	cumple
	Participantes en que se cumple				3
Frente a ASR	E1M1_R2	0	sin ventanas de este tipo		
	E1M2_R2	17	−80,4	−133	no cumple
	E2E3S7	10	−8,06	−5,17	no cumple
	E2E3S8	10	−55,7	−76,5	no cumple
	E2E3S9	10	−48,7	−68,4	no cumple
	Participantes en que se cumple				0
Frente al AF por bloques	E1M1_R2	0	sin ventanas de este tipo		
	E1M2_R2	17	−13,9	−49,7	no cumple
	E2E3S7	10	−4,27	−0,40	no cumple
	E2E3S8	10	1,13	9,15	cumple
	E2E3S9	10	2,43	14,2	cumple
	Participantes en que se cumple				2
Frente a la CCA con rezagos	E1M1_R2	0	sin ventanas de este tipo		
	E1M2_R2	17	3,62	−28,0	no cumple
	E2E3S7	10	−3,48	−3,94	no cumple
	E2E3S8	10	0,05	7,30	empate
	E2E3S9	10	0,44	−3,27	no cumple
	Participantes en que se cumple				0

Tabla D.3: Ventanas de masticación independientes del diseño: criterio 1 frente a cada comparador y requisito previo de iCanClean, calculados solo con esas ventanas, con la configuración elegida para cada participante. Ventanas: número de esas ventanas en cada participante, 47 en total. Diferencias en puntos porcentuales, de iCanClean menos el comparador o la identidad, con los márgenes de la Tabla 5.11. El requisito previo considera solo $\Delta\mathrm{RMS}_{\mathrm{mast}}$ (—). La última fila de cada comparación indica en cuántos de los 4 participantes con esas ventanas se cumple. Este análisis no asigna una etiqueta.

en $\Delta\mathrm{RMS}_{\mathrm{mast}}$); en E2E3S7, la del criterio 2 frente al AF pasó a cumplida (0,50 dB en la desviación máxima por banda), y frente a ASR, a empate.

Sin la última ventana evaluada de las piezas cuya última ventana de iCanClean se extiende hacia la guarda final, el requisito previo y la comparación con el AF en el criterio 1 cambiaron del mismo modo en E1M1_R2 (diferencia de 0 en el primero), y en E1M2_R2 la comparación con la CCA en el criterio 2 pasó a empate (0,0059 en la desviación de $Q_{\mathrm{RMS,reposo}}$).

Con el promedio de referencia calculado sin los canales de gran amplitud, en E1M2_R2, el único participante con canales marcados, ninguna comparación cambió de resultado, pero las diferencias de iCanClean en 20–45 Hz disminuyeron en valor absoluto, sin desaparecer: de −41,8 a −15,4 puntos porcentuales frente al AF por bloques y de −32,0 a −19,8 frente a la CCA con rezagos; frente a la identidad, la diferencia pasó de 5,57 a 4,51.

Método	Participante	S_{bal}	$\Delta\mathrm{RMS}_{\mathrm{mast}}$	Reducción en 20–45 Hz	R^2_{reposo}	$Q_{\mathrm{RMS,reposo}}$
ASR	E1M1_R2	71,9	52,6 %	4,17 %	91,2 %	0,957
	E1M2_R2	54,9	85,8 %	86,8 %	23,9 %	0,459
	E2E3S7	54,7	9,45 %	9,88 %	100,0 %	1,00
	E2E3S8	59,0	71,3 %	85,2 %	46,8 %	0,716
	E2E3S9	56,4	71,0 %	81,9 %	41,8 %	0,597
	Media	59,4	58,0 %	53,6 %	60,7 %	0,746
AF por bloques	E1M1_R2	52,8	6,64 %	0,27 %	98,9 %	0,988
	E1M2_R2	58,3	18,8 %	1,02 %	97,8 %	0,974
	E2E3S7	50,5	4,71 %	0,47 %	96,3 %	0,964
	E2E3S8	57,0	14,6 %	0,49 %	99,4 %	0,990
	E2E3S9	54,7	13,4 %	1,88 %	96,0 %	0,992
	Media	54,7	11,6 %	0,82 %	97,7 %	0,981
CCA con rezagos	E1M1_R2	50,7	1,69 %	−5,41 %	99,8 %	0,998
	E1M2_R2	50,4	2,93 %	−9,00 %	97,8 %	0,985
	E2E3S7	49,9	2,76 %	4,74 %	97,1 %	0,979
	E2E3S8	58,6	17,1 %	2,62 %	100,0 %	0,999
	E2E3S9	54,9	9,97 %	14,8 %	99,9 %	0,999
	Media	52,9	6,89 %	1,55 %	98,9 %	0,992
iCanClean	E1M1_R2	50,0	0 %	0 %	100,0 %	1,00
	E1M2_R2	52,8	5,64 %	−41,1 %	100,0 %	1,00
	E2E3S7	50,2	1,06 %	2,78 %	99,4 %	0,999
	E2E3S8	58,6	17,1 %	6,49 %	100,0 %	1,00
	E2E3S9	55,3	10,7 %	11,5 %	99,9 %	0,999
	Media	53,4	6,90 %	−4,06 %	99,9 %	1,00

Tabla D.4: Cercanía a la reserva de calibración: métricas de cada método, con la configuración elegida para cada participante en el análisis principal, sin la primera ventana evaluada de cada pieza que contiene una reserva, en masticación y en reposo; valor de cada participante y media. Sin esa ventana quedan 2, 53, 22, 22 y 22 ventanas de masticación, y 7, 10, 10, 10 y 10 de reposo, en E1M1_R2, E1M2_R2, E2E3S7, E2E3S8 y E2E3S9, respectivamente. Valores derivados de las métricas por ventana y de las particiones de la evaluación (Apéndice E). R^2, $\Delta\mathrm{RMS}$ y la reducción de potencia en 20–45 Hz se expresan en porcentaje.

D.4. Duración de la ventana de iCanClean

Las Tablas D.6 y D.7 presentan, con 2 s y con 8 s, la configuración elegida, sus métricas y la fracción de ventanas con selección; con 4 s, la elegida fue la del análisis principal (Tablas 5.6 y 5.5). Con 8 s, las medias de masticación y de S_{bal} incluyen 4 participantes y las de reposo, 5; las de los 4 participantes con valor en todas las duraciones (E1M2_R2, E2E3S7, E2E3S8 y E2E3S9) usan las mismas ventanas.

Ninguna duración obtuvo el mayor S_{bal} en todos los participantes; en los participantes comunes, su media fue 51,2, 54,1 y 53,9 con 2, 4 y 8 s. Con 2 s, iCanClean obtuvo la mayor $\Delta\mathrm{RMS}_{\mathrm{mast}}$ media, pero la menor reducción media en 20–45 Hz, negativa, y modificó más el reposo. Con 8 s, en esos participantes, atenuó algo más en masticación que con 4 s (9,80 % frente a 8,83 %), con una reducción en 20–45 Hz menos negativa (−2,92 % frente a −4,59 %) y un R^2_{reposo} algo menor (98,1 % frente a 99,3 %), sin un S_{bal} medio mayor.

	Cercanía a la reserva		Última ventana de iCanClean		Canales de gran amplitud	
Criterio	Cumple	Veredicto	Cumple	Veredicto	Cumple	Veredicto
Frente a ASR						
Criterio 1	0	no se cumple	0	no se cumple	0	no se cumple
Criterio 2	4	se cumple	4	se cumple	4	se cumple
Frente al AF por bloques						
Criterio 1	1	no se cumple	1	no se cumple	1	no se cumple
Criterio 2	1	no se cumple	0	no se cumple	0	no se cumple
Frente a la CCA con rezagos						
Criterio 1	0	no se cumple	0	no se cumple	0	no se cumple
Criterio 2	1	no se cumple	0	no se cumple	1	no se cumple
Frente a la identidad						
Requisito previo	4	se cumple	4	se cumple	5	se cumple
Etiqueta	no respaldada		no respaldada		no respaldada	
Rango de etiquetas	no respaldada		no respaldada		no respaldada	

Tabla D.5: Criterios de la hipótesis en las sensibilidades de diseño que recalculan la etiqueta (Sección 4.5.10), con la configuración elegida para cada participante; los del análisis principal están en la Tabla 5.10. Cumple: número de participantes, de cinco, en que se cumple cada criterio.

			Masticación		Reposo		No evaluables	
Participante	Configuración	S_{bal}	$\Delta\mathrm{RMS}$	20–45 Hz	R^2	Q_{RMS}	Mast.	Reposo
Ventanas de 2 s (grilla principal)								
E1M1_R2	ICC-08	61,3	23,9 %	−77,0 %	98,7 %	0,994	0	0
E1M2_R2	ICC-08	50,7	7,39 %	−54,3 %	93,9 %	0,969	0	0
E2E3S7	ICC-07	44,4	6,85 %	9,10 %	81,9 %	0,904	0	0
E2E3S8	ICC-08	57,5	16,2 %	−3,10 %	98,7 %	0,993	0	0
E2E3S9	ICC-08	52,3	10,6 %	1,94 %	94,0 %	0,969	0	0
Media	—	53,2	13,0 %	−24,7 %	93,4 %	0,966	0	0
Ventanas de 8 s (fuera de la grilla principal)								
E1M1_R2	ICC-W8-05	—	—	—	100,0 %	1,00	3	0
E1M2_R2	ICC-W8-07	52,5	4,96 %	−37,4 %	100,0 %	1,00	0	0
E2E3S7	ICC-W8-05	48,8	5,39 %	8,07 %	92,3 %	0,960	0	0
E2E3S8	ICC-W8-07	58,9	17,7 %	7,11 %	100,0 %	1,00	0	0
E2E3S9	ICC-W8-05	55,6	11,1 %	10,6 %	100,0 %	1,00	0	0
Media	—	53,9	9,80 %	−2,92 %	98,5 %	0,992	0,6	0

Tabla D.6: iCanClean con ventanas de 2 s y de 8 s: configuración elegida para cada participante entre las de cada duración, con los otros cuatro participantes, y sus métricas en el participante evaluado; con 4 s, son las del análisis principal (Tabla 5.6). Las configuraciones de 8 s se identifican en la Tabla B.3. Las métricas se calculan solo en las ventanas evaluables; las no evaluables son las que ninguna ventana completa de iCanClean de esa duración cubre. Con 8 s, E1M1_R2 no tiene ventanas evaluables en masticación, de modo que no tiene S_{bal} ni métricas de esa condición y queda fuera de sus medias. $\Delta\mathrm{RMS}$, la reducción de potencia en 20–45 Hz y R^2 se expresan en porcentaje.

	Todas las ventanas		Ventanas que cubre el control (a)			
	Propias		Propias		Otro participante	
Participante	Mast.	Reposo	Mast.	Reposo	Mast.	Reposo
Ventanas de 2 s (grilla principal)						
E1M1_R2	1,00	0,625	1,00	0,625	0,333	0,750
E1M2_R2	1,00	0,909	1,00	0,909	0,460	1,00
E2E3S7	1,00	0,909	1,00	0,909	0,348	0,545
E2E3S8	1,00	0,636	1,00	0,636	0,652	0,909
E2E3S9	1,00	1,00	1,00	1,00	1,00	1,00
Media	1,00	0,816	—	—	0,559	0,841
Ventanas de 8 s (fuera de la grilla principal)						
E1M1_R2	—	0	—	0	—	0
E1M2_R2	0,786	0	0,769	0	0	0
E2E3S7	0,833	0,333	0,833	0,333	0	0
E2E3S8	1,00	0	1,00	0	0	0
E2E3S9	0,833	0	0,833	0	0,333	0
Media	0,863	0,0667	—	—	0,0833	0

			Propias	Otro participante	
Ventana	Condición	Ventanas de iCanClean	Mediana	Mediana	Percentil 95
2 s	Reposo	52	0,709	0,709	0,783
	Masticación	122	0,867	0,673	0,836
4 s	Reposo	28	0,505	0,512	0,606
	Masticación	64	0,781	0,478	0,651
8 s	Reposo	14	0,342	0,306	0,385
	Masticación	31	0,701	0,298	0,500

			Umbral θ de la grilla				LIT-ICC-2
Ventana	Condición	Referencias	0,20	0,35	0,50	0,65	0,85
2 s	Reposo	Propias	1,00	1,00	1,00	0,827	0,0385
		Otro participante	1,00	1,00	1,00	0,846	0
	Masticación	Propias	1,00	1,00	1,00	1,00	0,557
		Otro participante	1,00	1,00	0,984	0,574	0,0410
4 s	Reposo	Propias	1,00	1,00	0,536	0,0714	0
		Otro participante	1,00	0,964	0,536	0,0357	0
	Masticación	Propias	1,00	1,00	1,00	0,781	0,203
		Otro participante	1,00	0,906	0,438	0,0625	0
8 s	Reposo	Propias	1,00	0,500	0,0714	0	0
		Otro participante	1,00	0,214	0	0	0
	Masticación	Propias	1,00	1,00	0,935	0,677	0,0645
		Otro participante	0,935	0,290	0,0645	0	0

Tabla D.7: iCanClean según la duración de la ventana, con las referencias propias y con las del control (a). Arriba, fracción de ventanas con al menos una componente retirada, con las configuraciones de la Tabla D.6 (con 4 s, en la Tabla 5.5); las columnas del control (a) se calculan en las ventanas que este cubre, y su fracción propia, sin media; en las filas por participante, una raya indica que no hay ventanas completas de iCanClean. Al centro, ρ_1^2, que con las referencias ajenas indica el nivel de azar frente al que se sitúan los umbrales θ; abajo, fracción de ventanas con $\rho_1^2 \geq \theta$ para cada umbral de la grilla y el de LIT-ICC-2, que no depende de la fuente, $\boldsymbol{U}$ o $\boldsymbol{V}$. Al centro y abajo, las ventanas de todos los participantes forman un solo conjunto, sin ponderar por participante.

Con 2 s, iCanClean retiró componentes también con las referencias de otro participante, en reposo en una fracción de ventanas igual o mayor que con las propias en E1M1_R2, E1M2_R2, E2E3S8 y E2E3S9; con 4 s y con 8 s, esa fracción fue baja. Estas fracciones concuerdan con el nivel de azar de ρ_1^2 (Tabla D.7): el mayor umbral de la grilla quedó bajo su mediana con 2 s y por encima de su percentil 95 con 8 s.

E. Trazabilidad

La Tabla E.1 identifica cada archivo analizado (Sección 4.1) por los primeros 16 caracteres hexadecimales de su suma SHA-256, que el código de la evaluación comprueba al abrirlo.

Archivo	SHA-256 del `.set`	SHA-256 del `.fdt`
`E1M1_R2_raw_step0`	`dcaeb672cbe324ed`	`1af2b68cef095cd5`
`E1M2_R2_raw_step0`	`d1ea8848de8ad690`	`176c99977327731d`
`E2E3S7_CH`	`059a580d10c6436d`	`b30289bd9051e160`
`E2E3S7_RS`	`0fc0a3605882bda6`	`ebae9e9d876b96bb`
`E2E3S8_CH`	`a4d9b86cda3a44a4`	`74a60961531c126b`
`E2E3S8_RS`	`b71d6eebf5eff3ea`	`f70837d443229121`
`E2E3S9_CH`[a]	`ce99f9487d7cdaed`	`fd868b1bc900e5b0`
`E2E3S9_RS`[a]	`9a17444596849b19`	`9b138cce96b21d6b`

Componente	Versión
Python	3.12.14
NumPy	2.5.4
SciPy	1.18.1
MNE-Python [1]	1.13.2
Matplotlib	3.11.2
scikit-learn	1.9.1
Plataforma de la evaluación	Linux-6.18.54.1-microsoft-standard-WSL2-x86_64-with-glibc2.43
asrpy [68]	0.0.8, con las adaptaciones del Apéndice A.2
Software de referencia, no ejecutado	
iCanClean, implementación oficial para EEGLAB [60]	Revisión `2189822`, utilizada como referencia de la formulación
`clean_rawdata`, implementación de referencia de ASR para EEGLAB [67]	Revisión `13e2efb`, utilizada como referencia de la formulación

Componente	Prefijo de la suma SHA-256
Identificador de la evaluación	`20261015T073152-0300-2770`
Protocolo de evaluación	`8b55ce86713de9be`
Configuración de la evaluación: decisiones y parámetros	`08f32611072ea04f`
Código de la evaluación: preprocesamiento, métodos, selección, métricas y figuras	`7182c1a588c4254d`
Tablas de resultados por ventana, canal, participante y método	`4dd3b1ab98b1e34e`
Tabla de cifras del texto	`721290467179d5e0`

Tabla E.1: Identificación de la evaluación: archivos analizados ([a]alias del participante: E2s3s9); versiones del software utilizado y del software de referencia; prefijos de la suma del protocolo de evaluación, del código y de los resultados del Capítulo 5, precedidos del identificador de la evaluación. Las sumas de cada archivo se conservan junto con los resultados.

Las versiones del software utilizado son las del entorno de la evaluación, y la copia adaptada de asrpy forma parte del código identificado en la tabla. Como ASR es sensible a cambios mínimos de su calibración (Apéndice A.2), sus cifras solo se reproducen exactamente con la misma plataforma y las mismas versiones de las bibliotecas. La evaluación se ejecuta con 3 procesos en paralelo y un solo hilo de cálculo numérico por proceso; cada participante se procesa de forma independiente, y una prueba con señales sintéticas dio resultados idénticos con uno y con varios procesos.

En la Tabla E.1, la suma del código corresponde a un registro con la suma de cada archivo de código de la evaluación, y la de las tablas de resultados, a un índice con la suma de cada tabla, señal guardada y figura de la evaluación. Las cifras de resultados de la evaluación no se transcriben a mano: se insertan desde la tabla de cifras, generada automáticamente desde las tablas de resultados.

Algunas cantidades que prescribe el protocolo de evaluación se derivan de forma exacta de sus tablas: la resolución de cada selección (Tabla B.2) y las métricas del análisis de cercanía a la reserva (Apéndice D.3). El programa que las deriva aplica las mismas reglas y fórmulas que la evaluación y comprueba las sumas de las tablas que lee; los prefijos de su suma y del registro de sus entradas y salidas son `7c0be71ca457181f` y `a9561df8749b83ca`. Las tablas no permiten recalcular, en cambio, la estabilidad frente a ASR de las comparaciones del criterio 2 en E1M2_R2, E2E3S8 y E2E3S9 ni las métricas de la sensibilidad a los canales de gran amplitud, para las que se informa el cálculo de la propia evaluación.

Las duraciones y el número de bloques del patrón EMG compatible con masticación en los archivos `ch` de E1 no provienen de la evaluación, sino del resumen de la búsqueda del Apéndice G. El programa que genera la tabla de cifras comprueba la suma de ese resumen, la del programa de detección y la de su módulo, cuyos prefijos son `37dceea6961f6cda`, `b6ea503ed43eb8a9` y `3b92affaede8d0a7`, y verifica en el registro de ejecución de la búsqueda que ese programa escribió ese resumen.

Las figuras de resultados del Capítulo 5 se dibujan a partir de las tablas de la evaluación y de las señales guardadas, sin cambiar ningún valor ni sus unidades. El programa que las dibuja, con prefijo `57c599932ca115d2`, comprueba la suma de cada tabla antes de leerla, y sus pruebas verifican que los valores dibujados coinciden con los de las tablas.

F. Resultados de la variante extendida

La variante extendida propia (Sección 3.6) se evaluó con el procedimiento de los demás métodos. Sus predictores, con desplazamientos entre $-L$ y L, se construyen sobre la pieza completa, de modo que las guardas aportan contexto a las muestras evaluadas de los extremos.

Su S_{bal} medio no alcanzó la ventaja mínima de un punto sobre el de iCanClean que exige el criterio de la Sección 4.5.8, y la Tabla F.1 presenta sus métricas y sus controles de especificidad.

Las configuraciones elegidas usan la envolvente y la menor escala γ de la grilla. En todas las selecciones, las configuraciones con envolvente y esa escala quedaron empatadas a menos de 0,5 puntos del mayor puntaje medio, cualquiera fuera su λ y su rezago máximo, de modo que la elección la decidió el desempate (Tabla B.2).

			Reposo		Masticación	
Participante	Configuración	S_{bal}	R^2	Q_{RMS}	$\Delta\mathrm{RMS}$	Reducción en 20–45 Hz
E1M1_R2	VAR-19	56,0	99,0 %	0,995	13,0 %	13,0 %
E1M2_R2	VAR-19	49,4	94,8 %	1,03	4,02 %	8,66 %
E2E3S7	VAR-20	49,0	95,3 %	1,03	2,62 %	3,92 %
E2E3S8	VAR-20	57,8	97,8 %	1,01	17,8 %	10,6 %
E2E3S9	VAR-31	45,8	82,2 %	1,16	9,43 %	21,6 %
Media	—	51,6	93,8 %	1,04	9,37 %	11,5 %
Configuración adicional, sin selección						
Media sin E1M1_R2	HIS-VAR	47,9	87,4 %	1,11	8,30 %	9,20 %

	$\Delta\mathrm{RMS}_{\mathrm{mast}}$				Reducción en 20–45 Hz				Fracción de $\Delta\mathrm{RMS}_{\mathrm{mast}}$	
Participante	Propias	(a)	(a′)	(b)	Propias	(a)	(a′)	(b)	(a)	(a′)
E1M1_R2	13,0 %	−8,01 %	−11,6 %	0 %	13,0 %	−7,86 %	−2,08 %	0 %	−0,618	−0,894
E1M2_R2	4,02 %	−1,47 %	−4,88 %	0 %	8,66 %	−3,28 %	−1,43 %	0 %	−0,371	−1,23
E2E3S7	2,62 %	−28,1 %	−1,08 %	0 %	3,92 %	−27,8 %	−1,29 %	0 %	−10,7	−0,413
E2E3S8	17,8 %	−28,6 %	−5,95 %	0 %	10,6 %	−85,3 %	−3,74 %	0 %	−1,61	−0,334
E2E3S9	9,43 %	−11,4 %	−1,80 %	0 %	21,6 %	−6,50 %	−0,41 %	0 %	−1,21	−0,190
Media	9,37 %	−15,5 %	−5,06 %	0 %	11,5 %	−26,1 %	−1,79 %	0 %	−2,90	−0,612

	R^2_{reposo}				$Q_{\mathrm{RMS,reposo}}$			
Participante	Propias	(a)	(a′)	(b)	Propias	(a)	(a′)	(b)
E1M1_R2	99,0 %	98,2 %	98,5 %	100,0 %	0,995	1,01	1,01	1,00
E1M2_R2	94,8 %	93,0 %	91,8 %	100,0 %	1,03	1,03	1,06	1,00
E2E3S7	95,3 %	97,0 %	99,8 %	100,0 %	1,03	1,02	1,00	1,00
E2E3S8	97,8 %	93,2 %	85,4 %	100,0 %	1,01	1,04	1,16	1,00
E2E3S9	82,2 %	88,5 %	97,7 %	100,0 %	1,16	1,08	1,02	1,00
Media	93,8 %	94,0 %	94,6 %	100,0 %	1,04	1,04	1,05	1,00

Tabla F.1: Variante extendida: valor de cada participante, con la configuración elegida con los otros cuatro participantes, y media. Arriba, configuración elegida, su puntaje balanceado S_{bal} en el participante evaluado y sus métricas principales; los identificadores remiten a la Tabla B.3, y la última fila corresponde a HIS-VAR (Tabla B.4), aplicada sin selección y resumida sin E1M1_R2. Al centro y abajo, controles de especificidad en masticación y en reposo, con las referencias propias y con los controles (a), (a′) y (b) de la Sección 4.5.7, y fracción de la atenuación que se mantiene, calculada en las mismas ventanas; los controles (a) y (a′) se evalúan en las ventanas que cubren, que solo en E1M2_R2 no son todas. R^2, $\Delta\mathrm{RMS}$ y la reducción en 20–45 Hz, en porcentaje.

El S_{bal} de la variante quedó bajo el de la identidad en E1M2_R2, E2E3S7 y E2E3S9. En E1M2_R2, E2E3S7, E2E3S8 y E2E3S9, $Q_{\mathrm{RMS,reposo}}$ fue mayor que uno, es decir, la variante aumentó la amplitud de la señal en reposo; el mayor aumento correspondió a E2E3S9, que tuvo también el menor R^2_{reposo}. Como R^2_{reposo} no depende de la escala, S_{bal} no penaliza ese aumento (Sección 4.5.6).

En E1M2_R2, $\Delta\mathrm{RMS}_{\mathrm{mast}}$ con las referencias propias fue 3,97 % en las ventanas que cubren los controles. Con las referencias de otro participante, con la calibración propia (a) o con una nueva (a′), la variante no atenuó la señal en masticación en ningún participante: $\Delta\mathrm{RMS}_{\mathrm{mast}}$ fue negativa y, con

(a), también lo fue la reducción en 20–45 Hz. Por ello, las fracciones de la Tabla F.1 son negativas; la media con (a) está dominada por E2E3S7, que tiene la menor reducción con las referencias propias. En la banda delta, la potencia en masticación disminuyó en todos los participantes con las referencias propias, entre $-2{,}55$ y $-0{,}25$ dB, y aumentó en todos con (a) y con (a′). En reposo, con (a), $Q_{\mathrm{RMS,reposo}}$ fue mayor que uno en todos los participantes. Así, la atenuación que la variante produce con las referencias propias no se mantuvo con las de otro participante.

Con referencias nulas (b), la salida fue idéntica a la entrada, bit a bit, en 13 de las 13 piezas comprobadas, de modo que el control verifica la implementación.

G. Búsqueda de masticación con el EMG en registros no evaluados

Fuera de la evaluación, y sin modificarla, se buscó masticación en registros de E1 no evaluados, con el EMG de los maseteros como única señal. Sus resultados describen un patrón EMG compatible con masticación, no la conducta, y se informan solo como limitación (Sección 5.8.5).

La regla de detección exige, en la envolvente de las derivaciones bipolares de los maseteros, ráfagas moduladas, simultáneas en ambos lados y rítmicas, con un nivel superior al del reposo. Sus umbrales se fijan con las reservas de calibración de los participantes evaluados, sin usar los archivos en que se aplica, y su especificidad frente a otras actividades orofaciales no se estimó.

En E1M3_R1, la regla no encontró masticación en el archivo `raw` ni en el de reposo: su único candidato en el primero era, a la inspección, un transitorio de un electrodo. En el archivo `ch` encontró candidatos, con ráfagas rítmicas y bilaterales en los primeros bloques; en los siguientes, cuyos electrodos inferiores presentaban ruido, derivas o valores cercanos a la saturación, un análisis exploratorio con los superiores, ajeno a la regla, sugiere el mismo patrón. Como estos candidatos no confirman la conducta, E1M3_R1 no se incorpora a la evaluación.

En E1M1_R2 y E1M2_R2, según los metadatos, los intervalos de masticación anotados son contiguos a bloques de aproximadamente un minuto de su archivo `ch`, que la evaluación no usa, de los que los separa un corte del archivo `raw`; en E1M1_R2, ese corte divide su intervalo de masticación (Tabla 4.1). Como la contigüidad se dedujo de los metadatos, y no de la señal de las piezas de evaluación, se desconoce si la masticación continúa a través de los cortes y forma parte del mismo episodio.

En esos archivos `ch`, la regla detectó el patrón, con certeza alta, en los 8 bloques de E1M1_R2, es decir, en todos, y en los 4 de la primera mitad del archivo de E1M2_R2. En los demás bloques de E1M2_R2 fallaron la sincronía y el ritmo del lado derecho, cuyo electrodo superior cambió de calidad; los archivos `rs` no se examinaron. En esos bloques, los intervalos candidatos ocuparon 424,5 s en E1M1_R2 y 148,0 s en E1M2_R2 en la estimación más ajustada, entre los centros de la primera y de la última ventana de análisis de cada candidato, y 456,0 y 197,5 s, respectivamente, como cota superior, con esas ventanas completas.

Estos bloques no forman parte de la evaluación, porque no tienen intervalos anotados ni una partición en calibración y evaluación. Para incorporarlos, habría que anotarlos, confirmar los

candidatos de E1M3_R1, establecer con quienes realizaron los registros qué marcan los códigos de evento y fijar sus particiones, sus reservas y su reposo antes de examinar en ellos los resultados de la limpieza; en E1M3_R1, la masticación y el reposo podrían emparejarse a partir de archivos distintos, como en E2. También convendría comprobar con la señal la continuidad de la masticación a través de los cortes, en un análisis separado, porque las piezas contiguas pertenecen a la evaluación; resolver la calidad del electrodo afectado en E1M2_R2; y estimar la especificidad de la regla.

H. Declaración de uso de herramientas de inteligencia artificial

En el desarrollo de esta tesis se utilizaron herramientas de inteligencia artificial generativa con el alcance que se describe a continuación.

- **Anotación de los intervalos.** Diez de los doce intervalos de reposo y masticación descritos en la Sección 4.1 se propusieron mediante una inspección asistida con OpenAI Codex, y su condición y extensión se ratificaron visualmente. Los dos intervalos restantes, correspondientes a E1M1_R2, provienen de la anotación manual original.
- **Versión previa del protocolo, código y versiones previas del texto.** En octubre de 2026 se utilizó OpenAI Codex como asistencia en la elaboración de una versión previa del protocolo de evaluación y en la obtención de sus resultados y figuras, que no son los que se informan en esta tesis. También se utilizó en la programación del código de análisis, parte del cual se reutiliza en la evaluación descrita en el Capítulo 4, y en la redacción de versiones previas de este documento.
- **Reconstrucción del texto, protocolo de evaluación y figuras.** En octubre de 2026 se utilizó Claude Code, de Anthropic, como asistencia en la reconstrucción del texto a partir del manuscrito original, en la verificación de las referencias bibliográficas y en la revisión de la redacción y de la forma del documento. Con esa misma asistencia se elaboró la propuesta detallada del protocolo de evaluación descrito en el Capítulo 4, sujeta a la aprobación del director, que comprende la partición en segmentos de calibración y de evaluación, la escala de los parámetros expresados en muestras, el procedimiento de selección, los criterios para examinar la hipótesis, la prueba semisintética y la definición operativa de la variante extendida, a partir de la regresión ridge con referencias rezagadas y del uso de la envolvente en el trabajo de la autora. También se programaron con esa asistencia los scripts que generan las figuras del Capítulo 1 y se elaboraron o modificaron los diagramas de los métodos y del procedimiento.

Las herramientas de inteligencia artificial no se consideran autoras de este trabajo. La autora es responsable del contenido de la tesis, de los métodos, de los resultados y de su interpretación.

Bibliografía

[1] A. Gramfort, M. Luessi, E. Larson, D. A. Engemann, D. Strohmeier, C. Brodbeck, R. Goj, M. Jas, T. Brooks, L. Parkkonen, and M. Hämäläinen, "MEG and EEG data analysis with MNE-Python," *Frontiers in Neuroscience*, vol. 7, p. 267, 2013.

[2] G. Tamburro, P. Fiedler, D. Stone, J. Haueisen, and S. Comani, "A new ICA-based fingerprint method for the automatic removal of physiological artifacts from EEG recordings," *PeerJ*, vol. 6, p. e4380, 2018.

[3] W. De Clercq, A. Vergult, B. Vanrumste, W. Van Paesschen, and S. Van Huffel, "Canonical correlation analysis applied to remove muscle artifacts from the electroencephalogram," *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 12, pp. 2583–2587, 2006.

[4] N. Kane, J. Acharya, S. Beniczky, L. Caboclo, S. Finnigan, P. W. Kaplan, H. Shibasaki, R. Pressler, and M. J. A. M. van Putten, "A revised glossary of terms most commonly used by clinical electroencephalographers and updated proposal for the report format of the EEG findings. Revision 2017," *Clinical Neurophysiology Practice*, vol. 2, pp. 170–185, 2017, Fe de erratas en Clin. Neurophysiol. Pract., vol. 4, p. 133, 2019, doi: 10.1016/j.cnp.2019.06.001.

[5] X. Jiang, G.-B. Bian, and Z. Tian, "Removal of artifacts from EEG signals: a review," *Sensors*, vol. 19, no. 5, p. 987, 2019.

[6] A. Vergult, W. De Clercq, A. Palmini, B. Vanrumste, P. Dupont, S. Van Huffel, and W. Van Paesschen, "Improving the interpretation of ictal scalp EEG: BSS–CCA algorithm for muscle artifact removal," *Epilepsia*, vol. 48, no. 5, pp. 950–958, 2007.

[7] S. Espinoza, W. El-Deredy, D. Moraga-Espinoza, and L. Morreal-Ortega, "Chewing phase - theta amplitude coupling facilitates working memory," bioRxiv preprint, version 1, Nov. 2025.

[8] S. L. Kappel, D. Looney, D. P. Mandic, and P. Kidmose, "Physiological artifacts in scalp EEG and ear-EEG," *BioMedical Engineering OnLine*, vol. 16, no. 1, p. 103, 2017.

[9] S. Espinoza, S. Cáceres, M. Salinas, D. Moraga-Espinoza, L. Morreal-Ortega, and W. El-Deredy, "Chewing modulates theta oscillation and functional connectivity of the frontocentral cortex in attention and working memory," *Scientific Reports*, vol. 15, no. 1, p. 43568, 2025.

[10] G. Krishnamoorthy, A. I. Narayana, and D. Balkrishanan, "Mastication as a tool to prevent cognitive dysfunctions," *Japanese Dental Science Review*, vol. 54, no. 4, pp. 169–173, 2018.

[11] R. J. Downey and D. P. Ferris, "The iCanClean algorithm: How to remove artifacts using reference noise recordings," arXiv preprint arXiv:2201.11798v1 [eess.SP], Jan. 2022.

[12] ——, "iCanClean removes motion, muscle, eye, and line-noise artifacts from phantom EEG," *Sensors*, vol. 23, no. 19, p. 8214, 2023.

[13] S. D. Muthukumaraswamy, "High-frequency brain activity and muscle artifacts in MEG/EEG: a review and recommendations," *Frontiers in Human Neuroscience*, vol. 7, p. 138, 2013.

[14] M. Seeck, L. Koessler, T. Bast, F. Leijten, C. Michel, C. Baumgartner, B. He, and S. Beniczky, "The standardized EEG electrode array of the IFCN," *Clinical Neurophysiology*, vol. 128, no. 10, pp. 2070–2077, 2017.

[15] V. Jurcak, D. Tsuzuki, and I. Dan, "10/20, 10/10, and 10/5 systems revisited: their validity as relative head-surface-based positioning systems," *NeuroImage*, vol. 34, no. 4, pp. 1600–1611, 2007.

[16] R. Oostenveld and P. Praamstra, "The five percent electrode system for high-resolution EEG and ERP measurements," *Clinical Neurophysiology*, vol. 112, no. 4, pp. 713–719, 2001.

[17] I. I. Goncharova, D. J. McFarland, T. M. Vaughan, and J. R. Wolpaw, "EMG contamination of EEG: spectral and topographical characteristics," *Clinical Neurophysiology*, vol. 114, no. 9, pp. 1580–1593, 2003.

[18] B. W. McMenamin, A. J. Shackman, L. L. Greischar, and R. J. Davidson, "Electromyogenic artifacts and electroencephalographic inferences revisited," *NeuroImage*, vol. 54, no. 1, pp. 4–9, 2011.

[19] E. M. Whitham, K. J. Pope, S. P. Fitzgibbon, T. Lewis, C. R. Clark, S. Loveless, M. Broberg, A. Wallace, D. DeLosAngeles, P. Lillie, A. Hardy, R. Fronsko, A. Pulbrook, and J. O. Willoughby, "Scalp electrical recording during paralysis: quantitative evidence that EEG frequencies above 20 Hz are contaminated by EMG," *Clinical Neurophysiology*, vol. 118, no. 8, pp. 1877–1888, 2007.

[20] C.-T. Lin, C.-S. Huang, W.-Y. Yang, A. K. Singh, C.-H. Chuang, and Y.-K. Wang, "Real-time EEG signal enhancement using canonical correlation analysis and Gaussian mixture clustering," *Journal of Healthcare Engineering*, vol. 2018, p. 5081258, 2018.

[21] J. A. Urigüen and B. Garcia-Zapirain, "EEG artifact removal—state-of-the-art and guidelines," *Journal of Neural Engineering*, vol. 12, no. 3, p. 031001, 2015.

[22] M. K. Islam, A. Rastegarnia, and Z. Yang, "Methods for artifact detection and removal from scalp EEG: a review," *Neurophysiologie Clinique/Clinical Neurophysiology*, vol. 46, no. 4–5, pp. 287–305, 2016.

[23] X. Chen, X. Xu, A. Liu, S. Lee, X. Chen, X. Zhang, M. J. McKeown, and Z. J. Wang, "Removal of muscle artifacts from the EEG: a review and recommendations," *IEEE Sensors Journal*, vol. 19, no. 14, pp. 5353–5368, 2019.

[24] H. Nolan, R. Whelan, and R. B. Reilly, "FASTER: fully automated statistical thresholding for EEG artifact rejection," *Journal of Neuroscience Methods*, vol. 192, no. 1, pp. 152–162, 2010.

[25] N. Bigdely-Shamlo, T. Mullen, C. Kothe, K.-M. Su, and K. A. Robbins, "The PREP pipeline: standardized preprocessing for large-scale EEG analysis," *Frontiers in Neuroinformatics*, vol. 9, p. 16, 2015.

[26] L. Wu, A. Liu, Z. J. Wang, R. K. Ward, and X. Chen, "Deep-learning based electroencephalogram denoising: a literature review," *Journal of Neural Engineering*, vol. 23, no. 4, p. 041003, Aug. 2026.

[27] H. Zhang, M. Zhao, C. Wei, D. Mantini, Z. Li, and Q. Liu, "EEGdenoiseNet: a benchmark dataset for deep learning solutions of EEG denoising," *Journal of Neural Engineering*, vol. 18, no. 5, p. 056057, Oct. 2021.

[28] H. Zhang, C. Wei, M. Zhao, Q. Liu, and H. Wu, "A novel convolutional neural network model to remove muscle artifacts from EEG," in *ICASSP 2021 – 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Toronto, ON, Canada, Jun. 2021, pp. 1265–1269.

[29] C.-H. Chuang, K.-Y. Chang, C.-S. Huang, and T.-P. Jung, "IC-U-Net: a U-Net-based denoising autoencoder using mixtures of independent components for automatic EEG artifact removal," *NeuroImage*, vol. 263, p. 119586, Nov. 2022.

[30] C. M. M. Queiroz, G. M. da Silva, S. Walter, L. B. Peres, L. M. D. Luiz, S. C. Costa, K. C. de Faria, A. A. Pereira, M. F. Vieira, A. M. Cabral, and A. d. O. Andrade, "Single channel approach for filtering electroencephalographic signals strongly contaminated with facial electromyography," *Frontiers in Computational Neuroscience*, vol. 16, p. 822987, 2022.

[31] G. Yilmaz, A. S. Budan, P. Ungan, B. Topkara, and K. S. Türker, "Facial muscle activity contaminates EEG signal at rest: evidence from frontalis and temporalis motor units," *Journal of Neural Engineering*, vol. 16, no. 6, p. 066029, 2019.

[32] T.-P. Jung, S. Makeig, C. Humphries, T.-W. Lee, M. J. McKeown, V. Iragui, and T. J. Sejnowski, "Removing electroencephalographic artifacts by blind source separation," *Psychophysiology*, vol. 37, no. 2, pp. 163–178, 2000.

[33] S. Makeig, A. J. Bell, T.-P. Jung, and T. J. Sejnowski, "Independent component analysis of electroencephalographic data," in *Advances in Neural Information Processing Systems*, D. Touretzky, M. C. Mozer, and M. Hasselmo, Eds., vol. 8. MIT Press, 1995, pp. 145–151. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/1995/hash/754dda4b1ba34c6fa89716b85d68532b-Abstract.html

[34] D. Gorjan, K. Gramann, K. De Pauw, and U. Marusic, "Removal of movement-induced EEG artifacts: current state of the art and guidelines," *Journal of Neural Engineering*, vol. 19, no. 1, p. 011004, 2022.

[35] L. Pion-Tonachini, K. Kreutz-Delgado, and S. Makeig, "ICLabel: an automated electroencephalographic independent component classifier, dataset, and website," *NeuroImage*, vol. 198, pp. 181–197, 2019.

[36] A. Delorme and S. Makeig, "EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," *Journal of Neuroscience Methods*, vol. 134, no. 1, pp. 9–21, 2004.

[37] A. Delorme, J. Palmer, J. Onton, R. Oostenveld, and S. Makeig, "Independent EEG sources are dipolar," *PLoS ONE*, vol. 7, no. 2, p. e30135, 2012.

[38] N. P. Castellanos and V. A. Makarov, "Recovering EEG brain signals: artifact suppression with wavelet enhanced independent component analysis," *Journal of Neuroscience Methods*, vol. 158, no. 2, pp. 300–312, Dec. 2006.

[39] H. Hotelling, "Relations between two sets of variates," *Biometrika*, vol. 28, no. 3–4, pp. 321–377, 1936.

[40] A. S. Janani, T. S. Grummett, T. W. Lewis, S. P. Fitzgibbon, E. M. Whitham, D. DelosAngeles, H. Bakhshayesh, J. O. Willoughby, and K. J. Pope, "Improved artefact removal from EEG using canonical correlation analysis and spectral slope," *Journal of Neuroscience Methods*, vol. 298, pp. 1–15, 2018.

[41] K. T. Sweeney, S. F. McLoone, and T. E. Ward, "The use of ensemble empirical mode decomposition with canonical correlation analysis as a novel artifact removal technique," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 1, pp. 97–105, 2013.

[42] X. Chen, X. Xu, A. Liu, M. J. McKeown, and Z. J. Wang, "The use of multivariate EMD and CCA for denoising muscle artifacts from few-channel EEG recordings," *IEEE Transactions on Instrumentation and Measurement*, vol. 67, no. 2, pp. 359–370, 2018.

[43] Y. Feng, Q. Liu, A. Liu, R. Qian, and X. Chen, "A novel SSA-CCA framework for muscle artifact removal from ambulatory EEG," *Virtual Reality & Intelligent Hardware*, vol. 4, no. 1, pp. 1–21, 2022.

[44] M. N. Anastasiadou, M. Christodoulakis, E. S. Papathanasiou, S. S. Papacostas, and G. D. Mitsis, "Unsupervised detection and removal of muscle artifacts from scalp EEG recordings using canonical correlation analysis, wavelets and random forests," *Clinical Neurophysiology*, vol. 128, no. 9, pp. 1755–1769, Sep. 2017.

[45] D. Safieddine, A. Kachenoura, L. Albera, G. Birot, A. Karfoul, A. Pasnicu, A. Biraben, F. Wendling, L. Senhadji, and I. Merlet, "Removal of muscle artifact from EEG data: comparison between stochastic (ICA and CCA) and deterministic (EMD and wavelet-based) approaches," *EURASIP Journal on Advances in Signal Processing*, vol. 2012, no. 1, p. 127, 2012.

[46] T. R. Mullen, C. A. E. Kothe, Y. M. Chi, A. Ojeda, T. Kerth, S. Makeig, T.-P. Jung, and G. Cauwenberghs, "Real-time neuroimaging and cognitive monitoring using wearable dry EEG," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 11, pp. 2553–2567, Nov. 2015.

[47] R. Somervail, J. Cataldi, A. M. Stephan, F. Siclari, and G. D. Iannetti, "Dusk2Dawn: an EEGLAB plugin for automatic cleaning of whole-night sleep electroencephalogram using artifact subspace reconstruction," *Sleep*, vol. 46, no. 12, p. zsad208, Dec. 2023.

[48] C.-Y. Chang, S.-H. Hsu, L. Pion-Tonachini, and T.-P. Jung, "Evaluation of artifact subspace reconstruction for automatic EEG artifact removal," in *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Honolulu, HI, Jul. 2018, pp. 1242–1245.

[49] S. Blum, N. S. J. Jacobsen, M. G. Bleichner, and S. Debener, "A Riemannian modification of artifact subspace reconstruction for EEG artifact handling," *Frontiers in Human Neuroscience*, vol. 13, p. 141, Apr. 2019.

[50] C.-Y. Chang, S.-H. Hsu, L. Pion-Tonachini, and T.-P. Jung, "Evaluation of artifact subspace reconstruction for automatic artifact components removal in multi-channel EEG recordings," *IEEE Transactions on Biomedical Engineering*, vol. 67, no. 4, pp. 1114–1121, Apr. 2020.

[51] M. Miyakoshi, "Artifact subspace reconstruction: a candidate for a dream solution for EEG studies, sleep or awake," *Sleep*, vol. 46, no. 12, p. zsad241, Dec. 2023.

[52] B. Widrow, J. R. Glover, Jr., J. M. McCool, J. Kaunitz, C. S. Williams, R. H. Hearn, J. R. Zeidler, E. Dong, Jr., and R. C. Goodlin, "Adaptive noise cancelling: principles and applications," *Proceedings of the IEEE*, vol. 63, no. 12, pp. 1692–1716, 1975.

[53] P. He, G. Wilson, and C. Russell, "Removal of ocular artifacts from electro-encephalogram by adaptive filtering," *Medical & Biological Engineering & Computing*, vol. 42, no. 3, pp. 407–412, 2004.

[54] P. He, G. Wilson, C. Russell, and M. Gerschutz, "Removal of ocular artifacts from the EEG: a comparison between time-domain regression method and adaptive filtering method using simulated data," *Medical & Biological Engineering & Computing*, vol. 45, no. 5, pp. 495–503, 2007.

[55] S. Mehrkanoon, M. Moghavvemi, and H. Fariborzi, “Real time ocular and facial muscle artifacts removal from EEG signals using LMS adaptive algorithm,” in *2007 International Conference on Intelligent and Advanced Systems*. Kuala Lumpur, Malaysia: IEEE, 2007, pp. 1245–1250.

[56] M. Kołodziej, A. Majkowski, and P. Wiszniewski, “Improved SSVEP classification through EEG artifact reduction using auxiliary sensors,” *Sensors*, vol. 26, no. 3, p. 917, 2026.

[57] A. D. Nordin, W. D. Hairston, and D. P. Ferris, “Dual-electrode motion artifact cancellation for mobile electroencephalography,” *Journal of Neural Engineering*, vol. 15, no. 5, p. 056024, 2018.

[58] C. B. Gonsisko, D. P. Ferris, and R. J. Downey, “iCanClean improves independent component analysis of mobile brain imaging with EEG,” *Sensors*, vol. 23, no. 2, p. 928, 2023.

[59] A. Studnicki, R. J. Downey, and D. P. Ferris, “Characterizing and removing artifacts using dual-layer EEG during table tennis,” *Sensors*, vol. 22, no. 15, p. 5867, 2022.

[60] R. J. Downey, “iCanClean EEGLAB plugin,” Código fuente MATLAB en GitHub, 2024, commit 2189822 (19 de noviembre de 2024); consultado el 14 de octubre de 2026. [Online]. Available: https://github.com/downeyryanj/iCanClean/tree/21898227823e2792c27717f144e4d4b9d6a5693c

[61] N. A. Jacobsen and D. P. Ferris, “Electrocortical activity correlated with locomotor adaptation during split-belt treadmill walking,” *The Journal of Physiology*, vol. 601, no. 17, pp. 3921–3944, 2023.

[62] J. Chmiel and A. Malinowska, “The neural correlates of chewing gum—a neuroimaging review of its effects on brain activity,” *Brain Sciences*, vol. 15, no. 6, p. 657, 2025.

[63] K. Sakamoto, H. Nakata, and R. Kakigi, “The effect of mastication on human cognitive processing: a study using event-related potentials,” *Clinical Neurophysiology*, vol. 120, no. 1, pp. 41–50, 2009.

[64] O. Sakata, K. Kohyama, T. Satake, and Y. Saito, “Extraction of alpha activities from an EEG obtained during gum chewing,” *IEEJ Transactions on Electrical and Electronic Engineering*, vol. 3, no. 3, pp. 324–333, 2008.

[65] BioSemi, “What is the function of the CMS and DRL electrodes. Ground, Reference or what?” Preguntas frecuentes del sitio web de BioSemi, consultado el 14 de octubre de 2026. [Online]. Available: https://www.biosemi.com/faq/cms&drl.htm

[66] C. A. E. Kothe and T.-P. Jung, “Artifact removal techniques with signal reconstruction,” U.S. Patent Application Publication 2016/0113587 A1, Apr. 28, 2016, Appl. No. 14/895,440; PCT/US2014/040770. [Online]. Available: https://patents.google.com/patent/US20160113587A1/en

[67] C. Kothe, M. Miyakoshi, and A. Delorme, “clean_rawdata,” Complemento de EEGLAB, versión 2.11, repositorio de GitHub sccn/clean_rawdata, 2025, commit 13e2efb44fcca4e6d64264438641e2fdfa9ec2f4 (10 de septiembre de 2025); consultado el 14 de octubre de 2026. [Online]. Available: https://github.com/sccn/clean_rawdata/tree/13e2efb44fcca4e6d64264438641e2fdfa9ec2f4

[68] D. Gütlin, “ASRpy: artifact subspace reconstruction in Python,” Paquete de Python, versión 0.0.8, PyPI, 2025, publicado el 19 de junio de 2025; consultado el 14 de octubre de 2026. [Online]. Available: https://pypi.org/project/asrpy/0.0.8/

[69] BioSemi, “Is skin preparation necessary when using BioSemi active electrodes?” Preguntas frecuentes del sitio web de BioSemi, consultado el 14 de octubre de 2026. [Online]. Available: https://www.biosemi.com/faq/skin_preparation.htm