POLITECNICO DI TORINO
Repository ISTITUZIONALE

Machine Learning Decoding of Dystonic States from EEG-LFP Signals in GPi-DBS: A Multimodal Case Study

Original
Machine Learning Decoding of Dystonic States from EEG-LFP Signals in GPi-DBS: A Multimodal Case Study / Barbi, E., Fino, E., Rossi, C., Di Rita, A., Peraio, S., Noris, A., Ciabattini, S., Lazzeri, M., Melani, F., Lenge, M., Giordano, F., Mesin, L.. - In: BIOENGINEERING. - ISSN 2306-5354. - 13:11(2026). [10.3390/bioengineering13111313]

Availability:
This version is available at: 11583/3017901 since: 2026-11-13T17:34:02Z

Publisher:
MDPI

Published
DOI:10.3390/bioengineering13111313

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

19 November 2026

bioengineering

MDPI

Article

Machine Learning Decoding of Dystonic States from EEG-LFP Signals in GPi-DBS: A Multimodal Case Study

Elisa Barbi [1], Edoardo Fino [2], Chiara Rossi [1], Andrea Di Rita [3], Simone Peraio [3], Alice Noris [3], Sara Ciabattini [1,4], Marta Lazzeri [3], Federico Melani [2,5], Matteo Lenge [1], Flavio Giordano [3,6] and Luca Mesin [7,*]

1 Neuroscience and Human Genetics Department, Meyer Children's Hospital IRCCS, 50139 Florence, Italy; elisa.barbi@meyer.it (E.B.); chiara.rossi@meyer.it (C.R.); sara.ciabattini@meyer.it (S.C.); matteo.lenge@meyer.it (M.L.)
2 Clinical Neurophysiology and Neuropsychology Lab, Meyer Children's Hospital IRCCS, 50139 Florence, Italy; edoardo.fino@meyer.it (E.F.)
3 Epilepsy Surgery and Functional Neurosurgery Unit, Meyer Children's Hospital IRCCS, 50139 Florence, Italy; andrea.dirita@meyer.it (A.D.R.); simone.peraio@meyer.it (S.P.); alice.noris@meyer.it (A.N.); marta.lazzeri2@edu.unifi.it (M.L.); flavio.giordano@meyer.it (F.G.)
4 Neuroscience, Pharmacology and Child Health Department, University of Florence, 50139 Florence, Italy
5 Pediatric Neurology Unit, Meyer Children's Hospital IRCCS, 50139 Florence, Italy
6 Department NEUROFARBA, University of Florence, 50139 Florence, Italy
7 Department of Electronics and Telecommunications, Politecnico di Torino, 10129 Turin, Italy
* Correspondence: luca.mesin@polito.it

Abstract

Background: Programming Deep Brain Stimulation (DBS) for pediatric dystonia remains largely empirical, as no validated electrophysiological biomarker currently guides parameter selection, and most studies treat each signal modality separately rather than exploiting their joint structure. Methods: In a single pediatric patient with generalized dystonia and bilateral globus pallidus internus (GPi) DBS implantation, we synchronized electroencephalography (EEG), local field potential (LFP), and electromyography (EMG) recordings across seven sessions spanning stimulator-OFF and multiple stimulation intensities. Dystonic muscle activity labels were derived from a validated quantitative EMG pipeline. Regularized logistic regression classifiers, trained on EEG-only, LFP-only, and combined EEG-LFP feature sets, were evaluated under four validation schemes: within-OFF cross-validation, OFF-to-ON generalization, a mixed leave-one-recording-out (LORO) scheme, and generalization to unseen stimulation intensities. Results: The Combined model performed best in the mixed LORO scenario (balanced accuracy 0.74, sensitivity 0.78, specificity 0.70), inheriting a favorable trade-off from the complementary biases of EEG (sensitivity-dominated) and LFP (specificity-dominated). LFP-derived Hjorth mobility, the beta/gamma power ratio, EEG central alpha power, and central-temporal gamma coherence ranked among the most consistent candidate predictive features across scenarios. Generalization to unseen stimulation intensities preserved a similar, moderately degraded performance profile (balanced accuracy 0.66–0.70). Conclusions: These findings provide proof-of-concept evidence that multimodal EEG-LFP decoding can identify epochs of dystonic muscle activity in paediatric dystonia.

Keywords: paediatric dystonia; Deep Brain Stimulation; globus pallidus internus; GPi-DBS; Local Field Potentials; electroencephalography; quantitative electromyography; multimodal machine learning; biomarker; adaptive DBS

Check for updates

Academic Editors: Yifan Zhao, Yannis Trakadis and Antonio Brunetti

Received: 29 September 2026
Revised: 20 October 2026
Accepted: 4 November 2026
Published: 6 November 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Among hyperkinetic movement disorders, dystonia is defined by the co-occurrence of abnormal postures and movements, sustained or intermittent, often combined with overflow activity in adjacent muscle groups [1–4]. Its nosology rests on two complementary axes: one describing the clinical phenomenology, the other the underlying cause. Etiologically, children differ considerably from adults [5]: most pediatric cases result from early basal ganglia damage or dysfunction within related circuits, principally due to perinatal injury culminating in cerebral palsy, as well as a heterogeneous group of genetic disorders. Regardless of cause, these conditions converge on a shared physiological mechanism, an altered signal transmission across the basal ganglia–cerebellum–thalamus–cortex loop [6,7]. In generalized, severe forms, pharmacological therapy often fails, leaving affected children with marked limitations in ambulation, speech, and daily functioning, and a substantial burden on quality of life [8,9].

For these medication-refractory cases, Deep Brain Stimulation (DBS) targeting the globus pallidus internus (GPi) has become a recognized surgical neurostimulation treatment [10]. Its benefits are more consistently reproducible in monogenic forms, whereas dystonia resulting from acquired perinatal injury or a hereditary degenerative disease tends to follow a less predictable course, with a therapeutic response that varies considerably from one patient to another [11]. Part of this inconsistency likely stems from differences in anatomical organization, compounded by an incomplete understanding of which stimulation parameters work best, and further exacerbated by the absence of established neurophysiological markers to guide this choice.

The pediatric setting adds further layers of complexity, as the hardware and programming logic of DBS were designed primarily for the adult brain; consequently, applying them to a nervous system that is still developing presents challenges, such as electrode migration during growth, that are relatively rare in clinical practice with adults [12]. Programming itself is harder in dystonia than in other indications of DBS; while Parkinson's disease tends to respond to stimulation in a fairly predictable and timely manner, the benefit in dystonia often takes longer to emerge and varies considerably from patient to patient, a direct consequence of how heterogeneous the pathophysiology and symptoms of the disorder can be [13]. This absence of a reliable, real-time signal to guide parameter selection has motivated growing interest in electrophysiology-driven programming strategies, since clinicians cannot feasibly explore the full combinatorial space of amplitude, frequency, and contact configurations by trial and error alone [14]. Without standardized protocols to fall back on, clinicians are largely left to program each child's device empirically, a gap that quantitative, objective methods are well positioned to fill [15].

One promising direction has been to look at Local Field Potentials (LFP) recorded directly from the GPi. In parallel, electroencephalography (EEG) offers a non-invasive view of cortical activity, recording from scalp electrodes the voltage fluctuations produced by the synchronous synaptic activity of large cortical neuronal populations [16,17]. Symptomatic periods have repeatedly been linked to excess oscillatory power at both ends of the spectrum, low frequencies (δ, θ) and high frequencies (γ) [18], while asymmetries in γ-band activity between hemispheres appear to track which limb is affected [19]. Similar spectral signatures have also been described at the cortical level, and there is emerging evidence that DBS itself reshapes the flow of information between deep and cortical structures [20], as captured by directed-connectivity measures such as conditional Transfer Entropy [21].

Recent work has refined this picture. Chronic recordings from the pallidum and sensorimotor cortex in isolated dystonia showed that low-frequency (3–12 Hz) oscillations at both sites track the severity of dystonic signs, that cortical low-frequency activity is reduced by therapeutic stimulation, and that stimulation can induce entrained cortical

gamma oscillations [22]. At the scalp level, pallidal DBS in cervical dystonia changes cortical low-frequency power, and suppression over motor areas relates to clinical improvement [23]. Pallidal and subthalamic low-frequency activity and their coupling have also been linked to symptom severity [24], and a recent systematic review concluded that low-frequency pallidal activity is the most consistent LFP finding, whereas validated biomarkers for adaptive stimulation are still lacking [25]. In children, chronic recordings with the Percept system showed that pallidal beta activity characterizes status dystonicus [26], and beta peaks dominate during voluntary movement in dystonic cerebral palsy [27].

However, taken together, these studies share a common limitation: rather than building models capable of exploiting the complete, joint structure of simultaneously recorded central and peripheral activity, they treat each signal and each frequency band separately, looking for statistical differences. This univariate approach is also difficult to translate into something practically useful, such as an automated, epoch-by-epoch detector of pathological activity, useful in the context of closed-loop or adaptive DBS systems [28]. Multivariate, machine-learning-based decoding of neural signals has already proven valuable for adaptive DBS in other movement disorders, most notably Parkinson's disease, but comparable approaches remain largely unexplored in pediatric dystonia [29]. A second limitation lies in the identification of dystonic episodes. Because most pediatric studies rely on clinical scales or retrospective video analyses, the temporal resolution and reproducibility of subsequent electrophysiological evaluations are inherently restricted. Consequently, there is a critical need for automated methods to enable precise labeling, for which electromyography (EMG) is a natural candidate.

To bridge these two gaps, we paired a validated, quantitative electromyography (qEMG)-based detector of dystonic muscle activity with a multimodal machine-learning framework. Using synchronized EEG and LFP recordings from a pediatric patient with generalized dystonia and bilateral GPi-DBS implantation, we extracted a broad set of multivariate features and their cross-modal interactions. These features were then used to train logistic regression classifiers capable of distinguishing dystonic from non-dystonic muscle activity. Specifically, this work aimed to: (i) determine whether EEG-only, LFP-only, and combined EEG-LFP feature sets could discriminate dystonic from non-dystonic muscle activity; (ii) test whether a classifier trained solely on stimulator-OFF data generalizes to stimulator-ON recordings, as a proxy for robustness to a clinically relevant condition unseen during training; (iii) evaluate performance under a mixed leave-one-recording-out scheme combining OFF and ON recordings, closer to a scenario of short-term, within-patient generalization; and (iv) identify which spectral, temporal, and cross-modal features contribute most consistently across these validation schemes. Taken together, this work is intended as a methodological proof-of-concept for a reproducible, quantitative pipeline that could support the discovery of electrophysiological correlates of dystonic activity, contributing to a deeper characterization of the disorder's underlying neurophysiology.

2. Materials and Methods

2.1. Patient and Multimodal Data Acquisition Protocol

Data were acquired from a 12-year-old patient with generalized dystonia secondary to neonatal hypoxic-ischemic encephalopathy (cerebral palsy), whose pre-operative phenotype comprised dystonic-dyskinetic tetraparesis, distal spasticity, intellectual disability, and focal epilepsy. Bilateral DBS electrodes targeting the GPi were implanted (B33015, Medtronic 110 Percept™ PC with BrainSense™ technology, Medtronic, Minneapolis, MN, USA; Figure 1), and recordings were performed eight months later, after an uncomplicated post-operative course. At the time of acquisition, the pharmacological regimen consisted of valproic acid, trihexyphenidyl, tetrabenazine, and baclofen. Throughout the session,

the patient was seated in a personal wheelchair with no explicit task imposed and remained predominantly at rest, with only occasional, minor spontaneous movements (e.g., turning to look for a parent).

Each electrode comprises four segments, two of which are directional with three contacts each, for a total of eight contacts per hemisphere, allowing both Ring (uniform activation of all contacts within a segment) and directional (focal) stimulation configurations (Figure 2). During a single post-operative follow-up session, seven consecutive five-minute recordings were acquired, in the chronological order of their IDs, each corresponding to a different combination of active contact configuration and stimulation intensity (stimulator OFF, Ring configuration, directional configuration proximal to the GPi, and directional configuration distal to the GPi, at increasing current intensities; see Figure 3). For each recording, we simultaneously acquired bilateral LFP from the GPi (Percept™ system, 250 Hz sampling rate), 21-channel surface EEG according to the international 10–20 system, surface EMG from the flexor and extensor muscles of both upper limbs (both recorded with a SystemPLUS Evolution amplifier, Micromed, Mogliano Veneto, Italy, at a sampling rate of 2048 Hz), and an additional Trigger channel, placed over the implanted pulse generator and used as a common temporal reference for multimodal synchronization.

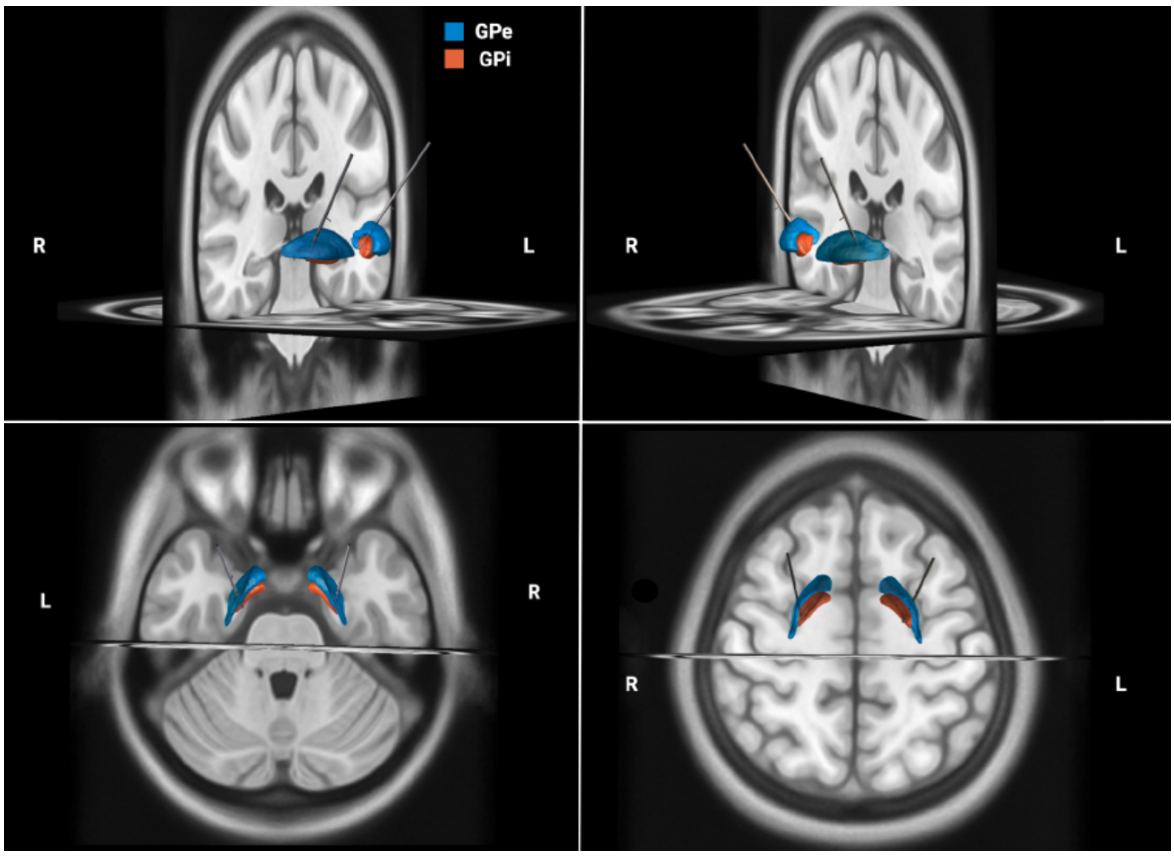


Figure 1. Post-operative three-dimensional reconstruction of the bilateral electrode trajectories relative to the GPe (blue) and GPi (orange), obtained with Lead-DBS [30].

Because the Ring configuration activates all contacts within a segment simultaneously, it produces a markedly different current-spread geometry and LFP artifact profile compared to the directional configurations used elsewhere in the protocol. Recording 2, the only session acquired with a non-directional Ring configuration, was therefore excluded from all quantitative analyses.

Consecutive 5-min blocks were separated by an interval of at least 20 min, used to export the telemetry data, adjust the stimulation parameters and allow the patient to settle. This interval was not intended as a clinical washout (see Section 4.1).



Figure 2. Schematic representation of the bilateral directional DBS electrodes implanted in the GPi. Each electrode comprises four levels: two non-directional ring contacts and two directional levels, each split into three equally spaced segments, allowing both Ring and directional (focal) stimulation configurations. Contacts are labeled according to the manufacturer's standard nomenclature (Medtronic SenSight): 0, 1a–1c, 2a–2c, 3 on the left electrode, and 8, 9a–9c, 10a–10c, 11 on the right electrode.

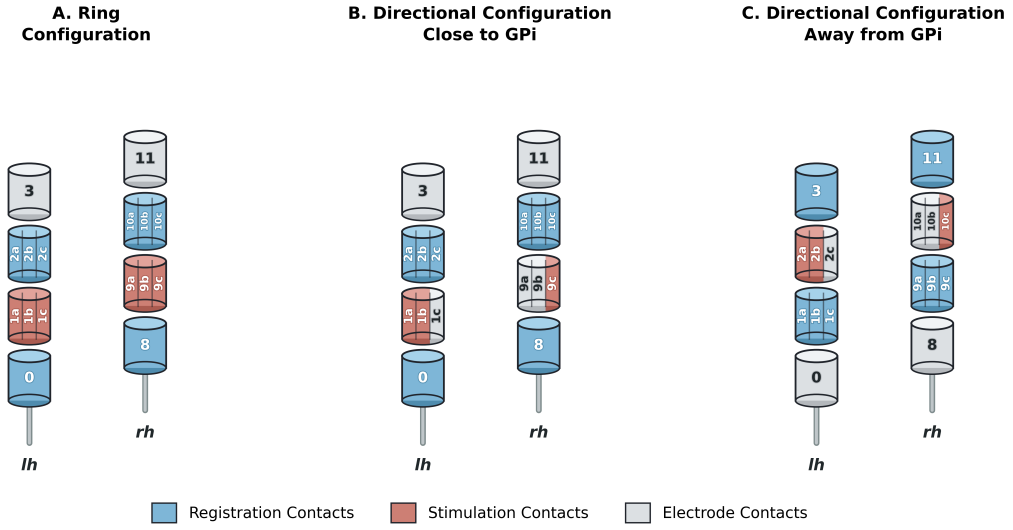


Figure 3. Schematic representation of the distinct stimulation and registration configurations applied during the recording protocol. (**A**) Ring configuration, with uniform activation across the directional segment. (**B**) Directional configuration proximal to the GPi. (**C**) Directional configuration distal to the GPi. Red cylinders indicate active stimulation contacts, blue cylinders denote registration contacts, and grey cylinders represent inactive contacts across both the left (lh) and right (rh) hemispheres.

2.2. Multimodal Signal Synchronization

We adapted the synchronization procedure established in our previous work on this recording setup [31]. LFP data were exported from the Percept™ system as JSON packets (62–63 samples each, 250 Hz); because the device occasionally drops or interrupts transmission, we scanned inter-packet timestamps and flagged any gap above the nominal 250 ms as a loss event, filling the corresponding span with NaN placeholders to keep the recordings temporally consistent ahead of alignment. In parallel, the higher-rate EEG/EMG streams

(2048 Hz) were downsampled to the common 250 Hz grid, using a fourth-order low-pass Butterworth filter (120 Hz cutoff) followed by a fourth-order notch filter around the 50 Hz mains frequency (49–51 Hz). Because the LFP, EEG and EMG acquisitions run on independent clocks, we introduced a set of short, clearly identifiable 2.4 mA stimulation pulses at the start and end of each session; these appear as sharp, synchronous deflections on both the Trigger channel and the stimulation-current trace and were used to register the three data streams onto a common time axis. Once the streams were aligned, the inserted placeholder segments and their corresponding intervals in the EEG/EMG recordings were discarded, together with a short margin at the onset and offset of each stimulation block, to avoid contaminating the subsequent analysis with transition artifacts.

Packet loss is reported in Supplementary Section S3.

2.3. Quantitative EMG-Based Detection of Dystonic Events

To obtain an objective, epoch-level reference labels for dystonic activity, where an epoch denotes a non-overlapping 1-s segment of the recording, we relied on the quantitative EMG (qEMG) pipeline previously developed and validated by our group [31], applied here to the four bilateral flexor and extensor channels. By definition, an epoch was classified as dystonic whenever it exhibited sustained, above-baseline activity co-occurring in both the flexor and the extensor of the same limb, reflecting agonist-antagonist co-contraction rather than a typical single-muscle movement pattern. To translate this definition into a detection algorithm, each channel was preprocessed independently: a fourth-order Butterworth high-pass IIR filter (10 Hz cutoff), applied in zero-phase mode removed the DC offset and the resulting signal was full-wave rectified.

A channel-specific reference level was then obtained by identifying the 4-s reference window with the lowest logarithmic energy across the entire recording, which was then used as a resting baseline against which all subsequent activity was compared. The logarithmic scale was adopted because energy is strictly positive and right-skewed; on this scale, fluctuations become approximately additive, so mean- and SD-based criteria are more meaningful, and epoch-to-epoch differences correspond to energy ratios, making both criteria independent of the absolute signal gain.

Working with the non-overlapping 1-s epochs (unrelated to the sliding reference window above), we computed for each epoch its logarithmic energy and marked it as potentially active if two conditions occurred simultaneously: its logarithmic energy exceeded the resting baseline by at least one standard deviation, and its rate of change (the absolute difference in logarithmic energy from the preceding epoch) remained below the 95th percentile of the recording. This latter constraint was included to filter out rapid, transient power spikes that accompany voluntary movement, as opposed to the more prolonged activation pattern expected during a dystonic episode.

Finally, to exclude isolated activations, only sequences of at least four active epochs were retained as candidate events; within these, the co-contraction criterion described above was applied to assign the final label dystonic, while the remaining epochs (active but lacking bilateral involvement of agonists and antagonists) were classified as non-dystonic motor activity.

Reference labels were established by two expert clinical neurophysiologists, who independently annotated the entire recordings, marking the epochs they judged to be dystonic without access to the detector output, and then resolved their disagreements by consensus. These consensus labels were used in all downstream analyses, and the automated detector was evaluated against them only afterwards, yielding a mean precision of 79.93 $\pm$ 0.17%, a recall of 90.09 $\pm$ 0.09%, and an F1-score of 83.28 $\pm$ 0.10%. A representative example of

the synchronized multimodal signals, together with the dystonic events identified by the qEMG pipeline, is shown in Figure 4.

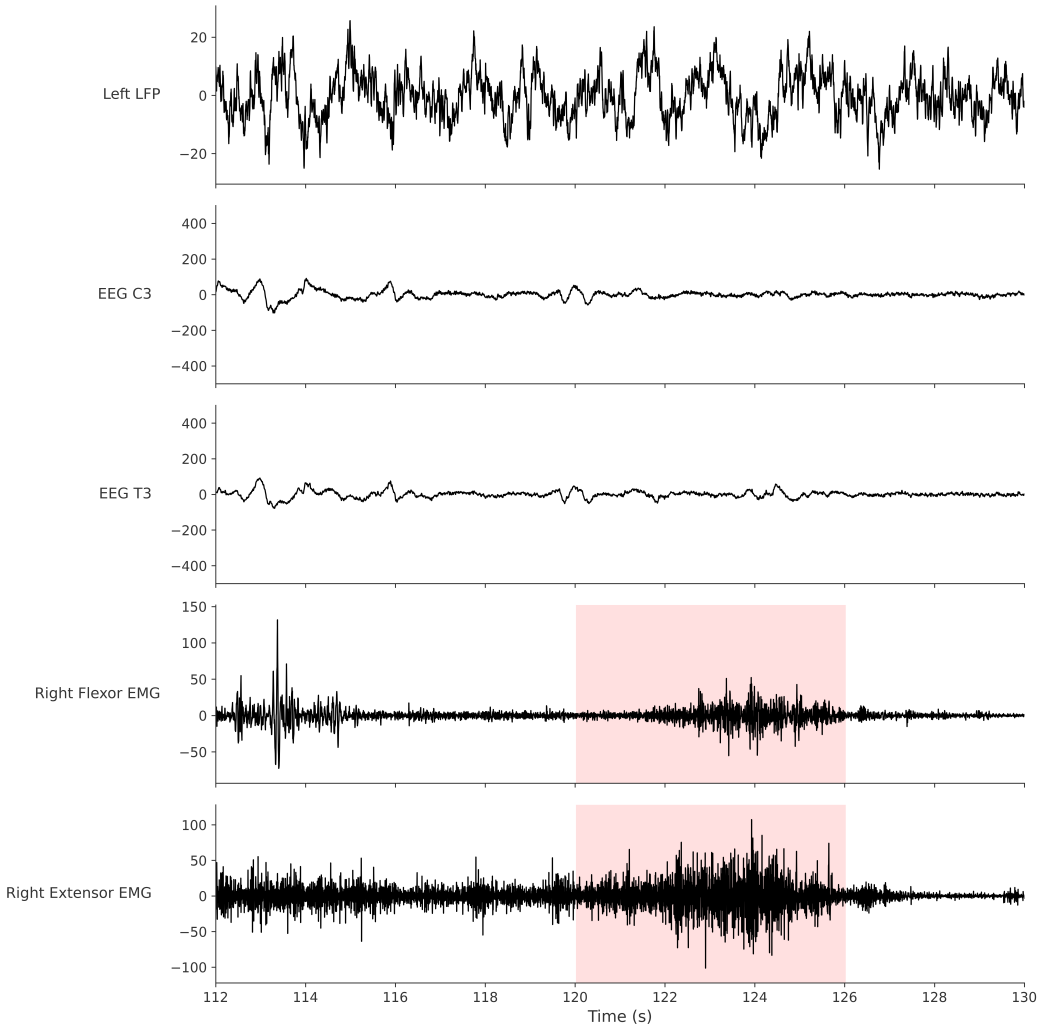


Figure 4. Example of synchronized multimodal signals and quantitative EMG-based labeling of dystonic muscle activity. The plot displays the LFP from the left GPi, ipsilateral EEG (C3, T3), and contralateral upper limb flexor and extensor EMG recordings. Shaded red areas highlight the dystonic muscle activity identified by the qEMG pipeline.

2.4. Preprocessing

For each recording (identified by a unique recording ID), the raw signal files were paired with the corresponding dystonic-event labels derived from the qEMG pipeline. Analyses were carried out separately for each hemisphere, combining the central and temporal EEG channels (C3/T3 on the left, C4/T4 on the right) with the ipsilateral LFP channel, while the dystonic-event label was taken from the contralateral clinical field.

This spatial restriction was clinically advised, and the channels were selected a priori, in agreement with the clinical neurologists, to limit the number of features while targeting key nodes of the dystonic network. C3/C4 overlie the sensorimotor cortex contralateral to the limb whose EMG provided the labels [32], and T3/T4 were added as adjacent lateral sites, since pallido-cortical coupling in dystonia has been described over both sensorimotor and temporal regions [33].

EEG signals were re-referenced to Cz, and all channels were visually inspected for bad or noisy segments. Independent component analysis (ICA) was then applied and the

resulting components were automatically classified with ICLabel; those identified as muscle, eye-movement or channel-noise components with a probability above 85% were removed.

Both EEG and LFP signals were filtered using a fourth-order Butterworth band-pass filter (0.5–45 Hz), followed by a standard 50 Hz IIR notch filter. For EEG channels only, we additionally applied two narrow-band second-order IIR notch filters (29 Hz and 37.8 Hz) to attenuate stimulation-related artifacts empirically identified in the power spectra. The effect of preprocessing on the spectra is shown in Supplementary Figure S4.

All filters were applied in zero-phase, forward-backward mode (`filtfilt`) to avoid phase distortion; this non-causal approach is appropriate for the present offline analysis, where edge effects are confined to the start and end of the signal. A causal implementation, required for real-time or online monitoring applications, would introduce a fixed group delay and would need to handle edge effects at the start of each epoch, while those at the end could not be resolved due to the lack of future samples.

Each filtered signal was then segmented into 1 s epochs, preserving consistency with the qEMG labels across all channels of a given segment.

Table 1 summarizes, for each recording and hemisphere, the active contacts, stimulation parameters, and the resulting epoch counts after preprocessing and qEMG-based labeling.

Table 1. Recording-level overview: stimulation parameters, active contacts, LFP sensing contacts, epoch counts, and event details per hemisphere. Recording ID corresponds to acquisition order.

Rec.	Hemi.	Total Epochs	Dystonic Epochs	Stimulation	Active Contacts	LFP Sensing Contacts
1	L	522	53	OFF	OFF	0/2a–c
1	R	522	220	OFF	OFF	8/10a–c
3	L	348	64	1.6 mA, 60 µs, 125 Hz	1a, 1b	0/2a–c
3	R	348	8	1.6 mA, 60 µs, 125 Hz	9c	8/10a–c
4	L	320	122	3.2 mA, 60 µs, 125 Hz	1a, 1b	0/2a–c
4	R	320	158	3.2 mA, 60 µs, 125 Hz	9c	8/10a–c
5	L	310	59	OFF	OFF	1a–c/3
5	R	310	95	OFF	OFF	9a–c/11
6	L	321	36	1.6 mA, 60 µs, 125 Hz	2a, 2b	1a–c/3
6	R	321	35	1.6 mA, 60 µs, 125 Hz	10c	9a–c/11
7	L	315	36	2.4 mA, 60 µs, 125 Hz	2a	1a–c/3
7	R	315	82	2.4 mA, 60 µs, 125 Hz	OFF	9a–c/11

2.5. Feature Extraction

We computed a set of spectral, temporal, and inter-channel features. Using power spectral density estimates obtained via Welch's method [34], we derived relative power in five canonical frequency bands (delta, 1–4 Hz; theta, 4–8 Hz; alpha, 8–13 Hz; beta, 13–30 Hz; gamma, 30–45 Hz), normalized to total power in the 0.5–45 Hz range, along with three band-power ratios (theta/beta, theta/alpha, beta/gamma).

Temporal dynamics within each epoch were characterized using the Hjorth mobility and complexity parameters [35], together with a set of time-domain statistics (skewness, kurtosis and zero-crossing rate). Furthermore, we assessed interchannel coupling using Welch-based coherence, computed for each frequency band, both within the EEG montage (between the central and temporal channels of the same hemisphere) and across modalities (between each EEG channel and the corresponding LFP channel).

Finally, to capture the temporal evolution of each feature, we additionally calculated first-order differences relative to the prior epoch (delta features), separately for each recording/hemisphere combination; the first epoch of each series, lacking a preceding value, was set to zero.

2.6. Classification Pipeline

Epoch-level classification (dystonic vs. non-dystonic) was performed using regularized logistic regression, comparing three input modalities:

- **EEG-only**: features derived from the EEG channels alone (including intra-EEG coherence);
- **LFP-only**: features derived from the LFP channel alone;
- **Combined**: the union of EEG, LFP, and EEG-LFP cross-modal features.

We evaluated the pipeline under four complementary scenarios. The first three were applied independently to all three modalities (EEG, LFP, Combined); the stimulator-OFF condition comprised recordings 1 and 5 (Table 1), while the stimulator-ON condition was restricted to the two directional recordings acquired at the lowest, most comparable current intensity (1.6 mA; recordings 3 and 6), in order to keep the stimulation amplitude consistent across the compared conditions. The fourth scenario instead assessed generalization across different stimulation intensities, as described below.

1. **Recording-level cross-validation restricted to the OFF condition.** Since only two OFF recordings were available, this scenario reduced to a two-fold, leave-one-recording-out scheme: the model was trained on one OFF recording and tested on the other, and the roles were then reversed. Because each training fold consisted of a single recording, hyperparameters and decision threshold could not be tuned with a recording-grouped inner cross-validation; the single training recording was instead partitioned into consecutive 10-epoch blocks, which served as surrogate groups for the inner cross-validation used to select both the regularization parameters and the decision threshold.
2. **OFF-to-ON generalization.** The model was trained and tuned on the full set of OFF epochs, using recording-grouped inner cross-validation, and then evaluated on the complete set of ON epochs. This scenario probes whether the model generalizes to a stimulation condition it never encountered during training.
3. **Mixed leave-one-recording-out (LORO).** A selected dataset, combining the OFF condition (sessions 1 and 5) and the 1.6 mA ON condition (sessions 3 and 6), was evaluated in a leave-one-recording-out scheme. At each iteration, a single session was held out for testing while the model was tuned via grouped inner cross-validation on the remaining data. Because the stimulation state of the test set was tracked throughout, performance could subsequently be broken down separately for the OFF and ON conditions, allowing us to verify whether the model generalized consistently across states.
4. **Cross-intensity generalization.** To determine whether the model maintains its predictive performance across different current intensities, the Combined classifier was trained using the dataset of OFF and 1.6 mA ON conditions (recordings 1, 3, 5, and 6). Evaluation was then performed on recordings 4 (3.2 mA) and 7 (2.4 mA), to serve as an independent test set with unseen stimulation parameters.

Within each fold of every scenario, the classifier was trained through the same three-step pipeline. First, we performed correlation filtering: pairwise Pearson correlation was computed across all candidate features using training data only, and for each pair exceeding $|r| > 0.85$, the feature less correlated with the target label was discarded; then, the remaining features were standardized via z-scoring. Finally, a logistic regression classifier with balanced class weights was trained using the liblinear solver under either an L1 or L2 penalty, with the regularization strength C selected from $\{0.01, 0.1, 1, 10\}$ via grid search; hyperparameters were tuned by optimizing balanced accuracy on an inner cross-validation restricted to the training portion of each outer fold, using up to 5 folds grouped by recording identifier. This number was dynamically capped by the available groups in the training

partition, resulting in 2 folds for both the OFF condition and OFF-to-ON generalization scenarios, and 3 folds for the mixed LORO scheme. Since dystonic muscle acitivity represent a minority class (Table 1), the decision threshold applied to predicted probabilities was not fixed at 0.5, but instead optimized for each fold and scenario using training data exclusively. Out-of-fold predictions were generated through the same inner cross-validation split used for hyperparameter tuning, and the threshold maximizing balanced accuracy was selected from a grid of equally spaced values between 0.1 and 0.9. This threshold was then applied, to the test set, ensuring that no test-set information contributed to its selection.

The resulting epoch-by-epoch binary predictions were further post-processed with a median filter to suppress isolated, temporally unstable misclassifications; a 5-epoch window was adopted for the main analysis. The sensitivity of the results to this choice was assessed in a dedicated comparison, in which raw predictions were computed only once per modality/scenario combination, avoiding repeated model training, and then smoothed with window (w) sizes of 1 (no smoothing), 3, 5, and 7 epochs, all evaluation metrics being recomputed for each window. To assess whether smoothing produced a statistically meaningful change in performance, we additionally performed a cluster bootstrap over recordings (2000 iterations): at each iteration, recording identifiers were resampled with replacement, and balanced accuracy, sensitivity, specificity, and pooled event-level detection rate were recomputed for every smoothing window on the resampled data. For each window, the resulting bootstrap distribution was compared to that of the raw (unsmoothed, $w = 1$) predictions in a paired fashion, and a 95% confidence interval for the difference was obtained from the empirical bootstrap distribution; a window was considered significantly different from raw when this interval excluded zero.

In every scenario, we explicitly verified at each fold that no recording appeared in both the training and the test set. To prevent leakage, all epochs from a given recording were constrained to the same fold throughout, using stratified group k-fold cross-validation grouped by recording identifier whenever the number of available recordings allowed it.

Classifier performance was quantified through balanced accuracy, sensitivity, and specificity, computed on both raw and smoothed predictions; sensitivity and specificity were obtained from the confusion matrix at the decision threshold selected as described above. All metrics are reported both as mean $\pm$ standard deviation across folds and as pooled values, the latter obtained by concatenating the out-of-fold predictions across all folds before computing each metric once on the full set.

Beyond these epoch-level metrics, we introduced an event-level evaluation to capture performance at the scale of individual episodes rather than single epochs. An event was defined as a run of consecutive epochs sharing a positive label, whether true or predicted. For each recording/hemisphere, we computed the number of true dystonic events together with the corresponding detection rate (the proportion of true events overlapping with at least one correctly predicted positive epoch), the number of false alarms (predicted events with no overlap with any true event) and the mean detection delay, defined as the number of epochs elapsing between the onset of a true event and the first epoch within it correctly flagged as positive.

To test whether performance exceeded chance, we compared the observed balanced accuracies with a null distribution obtained by label permutation. In each of 1000 permutations, the label sequence of each recording/hemisphere was circularly shifted by a random amount (at least 30 epochs), which breaks the link between signals and labels while preserving the temporal structure of the labels. The whole pipeline was repeated on the shifted labels for each modality and for the three main scenarios (the cross-intensity scenario was not included). The p-value was the proportion of the 1000 permutations in which the balanced accuracy was equal to or higher than the observed one (one was added

to both numerator and denominator, so that p is never zero and cannot be lower than 0.001). The same procedure was used to compare the Combined and EEG-only models: in each permutation, we computed the difference in balanced accuracy between the two models, and the p-value was the proportion of permutations in which this difference was equal to or larger than the one observed. p-values were adjusted for multiple comparisons with the Holm method.

2.7. Feature Importance Analysis

Feature importance was estimated by aggregating logistic regression coefficients across cross-validation folds. Because the correlation filter could retain slightly different feature subsets in different folds, aggregation was performed only over the folds in which a given feature was actually retained. For each feature, we computed the mean coefficient and its standard deviation across these folds, the number of folds in which the feature was retained, and a sign-consistency index, defined as the fraction of folds in which the coefficient retained the same sign. Features were then ranked by the absolute value of their mean coefficient, and the top-ranked features (up to 15) were displayed as horizontal bar plots, with bars colored according to the sign of the effect (pro-dystonic vs. pro-non-dystonic) and error bars representing the between-fold standard deviation.

As a control for possible cranial muscle contamination of high-frequency EEG content, the full pipeline was additionally repeated after excluding all EEG gamma-band features (Supplementary Materials, Section S1).

Finally, to clarify what the LFP-derived features capture, we characterized the LFP spectra of dystonic and non-dystonic motor activity. The preprocessed LFP signals used for classification were segmented into 1-s epochs, and the power spectral density of each epoch was estimated with Welch's method and expressed in $\mu V^2/Hz$. Spectra were pooled across recording/hemisphere series and parameterized into aperiodic and periodic components with specparam (fit range 3–40 Hz, fixed aperiodic mode). The upper limit of the fit range was set below the 45 Hz cutoff of the band-pass filter to limit the influence of its roll-off. Full details of the spectral estimation and parameterization are reported in Supplementary Materials, Section S2.

3. Results

3.1. Classification Performance

Table 2 reports epoch-level balanced accuracy, sensitivity, and specificity for the three modalities (EEG, LFP, Combined) across the three evaluation scenarios, before and after median-filter smoothing.

Restricted to the OFF condition, all three modalities performed only modestly above chance (balanced accuracy 0.59–0.62 after smoothing), with LFP and Combined showing a marked gap between specificity and sensitivity, indicating a systematic tendency to under-detect dystonic muscle activity when both training and testing were restricted to the stimulator-OFF condition; EEG showed the same direction of imbalance, but less pronounced. Generalizing from OFF to ON, EEG improved substantially on every metric, suggesting that the alpha-power-centered EEG signature learned on OFF data transfers well to the stimulation-ON condition. LFP, by contrast, remained close to chance level, indicating that whatever LFP-derived signature discriminates dystonic activity within a single stimulation condition does not generalize across stimulation states. The Combined model occupied an intermediate position (balanced accuracy 0.63 → 0.64), closer to sensitivity-specificity balance (0.61 vs. 0.69 after smoothing) than either single modality alone.

The mixed leave-one-recording-out scenario yielded the best overall performance for every modality, and in particular for the Combined model, which reached a balanced

accuracy of 0.74 (smoothed) together with the highest sensitivity observed in any scenario, 0.78, and a specificity of 0.70. EEG alone achieved comparable sensitivity (0.80) but at the cost of a substantially lower specificity (0.55); LFP alone showed the opposite asymmetry, with specificity remaining high (0.79) while sensitivity dropped to 0.47. This pattern, with EEG biased toward sensitivity and LFP toward specificity, is compatible with the more balanced profile of the Combined model in this scenario. Its advantage, however, was limited to this scenario: in the Recording scenario EEG-only performed equally well (0.62 vs. 0.61 after smoothing), and in the OFF-to-ON scenario it performed better (0.70 vs. 0.64). A formal comparison against a permutation null, including a paired Combined-versus-EEG test, is reported below.

Table 2. Epoch-level classification performance for the Logistic Regression classifier (raw → smoothed, $w = 5$).

Scenario	Modality	Bal. Acc.	Sensitivity	Specificity
Recording (OFF/OFF)	EEG	0.59 → 0.62	0.53 → 0.54	0.65 → 0.70
	LFP	0.61 → 0.60	0.49 → 0.44	0.73 → 0.75
	Combined	0.62 → 0.61	0.51 → 0.47	0.73 → 0.74
OFF → ON	EEG	0.65 → 0.70	0.60 → 0.64	0.69 → 0.75
	LFP	0.57 → 0.56	0.54 → 0.51	0.60 → 0.61
	Combined	0.63 → 0.64	0.64 → 0.61	0.63 → 0.69
LORO Mixed	EEG	0.65 → 0.67	0.75 → 0.80	0.54 → 0.55
	LFP	0.63 → 0.63	0.51 → 0.47	0.75 → 0.79
	Combined	0.72 → 0.74	0.76 → 0.78	0.67 → 0.70

To test whether these accuracies were higher than expected by chance, we compared them with the accuracies obtained after permuting the labels (Table 3). Balanced accuracy exceeded the null for all modalities in all three scenarios, and all nine tests remained significant after Holm correction. In the mixed LORO scenario, the Combined model reached 0.74 against a null mean of 0.52 (95th percentile 0.60; $p = 0.001$), EEG 0.67 (null 0.54; $p = 0.001$) and LFP 0.63 (null 0.49; $p = 0.001$). In the OFF-to-ON scenario, EEG reached 0.70 (null 0.52), Combined 0.65 (null 0.50) and LFP 0.56 (null 0.46), all with $p = 0.001$. Performance in the Recording scenario was the closest to the null (0.60–0.62 vs. null means of 0.51–0.52; $p = 0.010$–0.016), consistent with the modest accuracy described above. In the paired comparison, the Combined model outperformed EEG-only in the mixed LORO scenario ($\Delta = +0.065$; null 95th percentile 0.046; nominal $p = 0.024$), but this difference did not survive correction for the three comparisons (Holm-adjusted $p = 0.072$) and should be considered exploratory. No advantage was found in the OFF-to-ON ($\Delta = -0.05$; $p = 0.72$) or Recording scenarios ($\Delta = -0.01$; $p = 0.57$).

At the adopted 5-epoch window, median-filter smoothing improved balanced accuracy most consistently for EEG (+0.02 to +0.05 across scenarios, largest in OFF-to-ON), had a negligible or slightly negative effect for LFP in the Recording and OFF-to-ON scenarios (-0.01), and a small positive effect for the Combined model in the OFF-to-ON and LORO scenarios ($\leq +0.02$), suggesting that the isolated misclassifications targeted by the filter are more prevalent in the EEG-derived predictions than in the LFP-derived ones. Across the full range of window sizes tested (1, 3, 5, and 7 epochs; Figure 5), these trends were generally monotonic, and no single window size was uniformly optimal across modalities and scenarios: EEG and Combined balanced accuracy continued to increase beyond $w = 5$ in some scenarios (e.g., EEG reaching +0.06 in OFF-to-ON, Combined reaching +0.03 in LORO at $w = 7$), while LFP balanced accuracy in the LORO scenario, essentially unchanged at $w = 5$, dropped below the raw value at $w = 7$. For the main analysis, a 5-epoch window was

therefore chosen, as it captured most of the achievable improvement in balanced accuracy for the EEG and the Combined model, while avoiding wider windows in which performance began to decay for the LFP (Table 2).

Table 3. Permutation-based null analysis of pooled smoothed balanced accuracy ($w = 5$; 1000 circular-shift label permutations). Obs.: observed value; Null mean and Null P95: mean and 95th percentile of the null distribution; p: one-sided empirical p-value, uncorrected (0.001 is the minimum attainable with 1000 permutations); p_{Holm}: Holm-adjusted p-value (across the nine modality-versus-null tests, and separately across the three paired comparisons). Δ: difference in balanced accuracy between the Combined and EEG-only models (Combined − EEG), evaluated within each permutation.

Scenario	Modality	Obs.	Null Mean	Null P95	p	p_{Holm}
Recording (OFF/OFF)	EEG	0.621	0.524	0.590	0.010	0.030
	LFP	0.596	0.511	0.570	0.016	0.030
	Combined	0.607	0.514	0.580	0.014	0.030
OFF → ON	EEG	0.697	0.525	0.598	0.001	0.009
	LFP	0.562	0.461	0.500	0.001	0.009
	Combined	0.647	0.496	0.559	0.001	0.009
LORO Mixed	EEG	0.673	0.542	0.607	0.001	0.009
	LFP	0.630	0.489	0.563	0.001	0.009
	Combined	0.737	0.517	0.596	0.001	0.009
Paired comparison (Combined − EEG)						
Recording (OFF/OFF)	Δ	−0.014	−0.010	0.040	0.572	1.000
OFF → ON	Δ	−0.050	−0.029	0.029	0.718	1.000
LORO Mixed	Δ	+0.065	−0.025	0.046	0.024	0.072

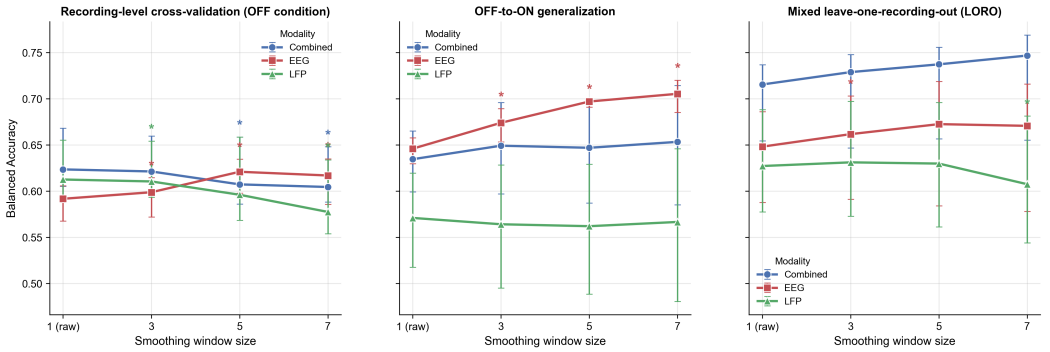


Figure 5. Balanced accuracy as a function of the median-filter smoothing window (1, 3, 5, 7 epochs), by modality, across the three main evaluation scenarios. Error bars denote 95% cluster-bootstrap confidence intervals over recordings. Asterisks mark smoothing windows for which balanced accuracy differed significantly from the raw predictions, based on a paired cluster-bootstrap comparison.

Performance computed separately for each hemisphere, recording and stimulation condition (Supplementary Section S5, Tables S4 and S5) showed no consistent difference between hemispheres, and the Combined model remained the best in both hemispheres in the mixed LORO scenario (balanced accuracy 0.69 left, 0.76 right). Performance was more heterogeneous across individual recording/hemisphere series, with balanced accuracy ranging from 0.49 to 0.76 for the Combined model in the mixed LORO scenario; the lowest values were observed in Recording 3, whose right hemisphere contained only 8 dystonic epochs.

3.2. Event-Level Detection Performance

Beyond epoch-level metrics, Table 4 reports event-level detection rate, number of false alarms, and mean detection delay for each scenario and modality (computed on smoothed predictions).

EEG consistently achieved the highest detection rate across all three scenarios (85.3%, 88.2%, and 94.1% for the Recording, OFF-to-ON, and LORO scenarios respectively), together

with the shortest detection delays throughout (0.54–0.78 epochs), but at the cost of the largest absolute number of false alarms in the Recording scenario. The Combined model matched EEG's detection rate in the OFF-to-ON scenario (88.2%) and approached it in the LORO scenario (90.2% vs. 94.1%), while incurring more false alarms and a longer mean detection delay than EEG. LFP alone showed the lowest detection rate throughout (55.9–70.6%) and the longest delays (1.57–2.31 epochs), but consistently the fewest false alarms among the three modalities, reinforcing the conservative behavior already observed at the epoch level.

Notably, in the Recording scenario the ranking of modalities by event-level detection rate does not fully mirror their epoch-level sensitivity: the Combined model has slightly higher smoothed epoch-level sensitivity than LFP alone (0.47 vs. 0.44, Table 2) but a lower event-level detection rate (50.0% vs. 55.9%, Table 4). This apparent discrepancy reflects the different unit of analysis between the two metrics: event-level detection only requires a single correctly flagged epoch within each true event, so a modality can achieve a higher overall epoch-level sensitivity while still missing entire events more often than a modality with more evenly distributed, if less numerous, correct detections. Given that the Recording scenario is based on only two OFF recordings and a two-fold leave-one-recording-out split, this divergence should be interpreted with caution and is best regarded as a further illustration of why both epoch-level and event-level metrics are reported.

Table 4. Event-level detection performance (smoothed predictions).

Scenario	Modality	Events Detected	Det. Rate	False Alarms	Mean Delay (Epochs)
Recording (OFF/OFF)	EEG	29/34	85.3%	53	0.78
	LFP	19/34	55.9%	22	2.31
	Combined	17/34	50.0%	26	2.10
OFF → ON	EEG	15/17	88.2%	47	0.64
	LFP	12/17	70.6%	56	2.02
	Combined	15/17	88.2%	52	1.92
LORO Mixed	EEG	48/51	94.1%	80	0.54
	LFP	34/51	66.7%	39	1.57
	Combined	46/51	90.2%	88	1.22

3.3. Feature Importance

To identify which signal features most consistently discriminated dystonic from non-dystonic muscle activity, we inspected the top-ranked logistic regression coefficients (aggregated across cross-validation folds) for each modality and scenario. Figure 6 reports the top-15 features for the Combined model in the mixed LORO scenario, which achieved the best overall performance and is therefore taken as the reference case; the same pattern of feature importance, detailed below, was also observed in the Recording and OFF-to-ON scenarios and in the single-modality models.

In the Combined model, LFP-derived Hjorth mobility was the strongest pro-dystonic predictor in all three evaluation scenarios (mean coefficient ≈1.1 in LORO Mixed, 1.5 in Recording, 1.5 in OFF-to-ON), while the LFP beta/gamma power ratio was symmetrically one of the two strongest pro-non-dystonic predictors throughout (≈−0.6 to −0.8).

On the EEG side, central relative alpha power was a pro-non-dystonic predictor in the Combined model across all three scenarios, while its first-order temporal derivative contributed with the opposite, pro-dystonic sign in every scenario, indicating that not only the instantaneous alpha power, but also its rate of change, carries independent discriminative information. Central-temporal gamma-band coherence and the EEG theta/beta

ratio completed the set of features that appeared, with the same sign (pro-dystonic and pro-non-dystonic, respectively), across all three Combined-model scenarios.

This pattern was not an artifact of the Combined feature set: when EEG and LFP were modeled separately, each single-modality model reproduced the same dominant features, with the same sign, in every scenario. In the LFP-only model, LFP-derived Hjorth mobility and the LFP beta/gamma power ratio were again the top two predictors (positive and negative, respectively) across the Recording, OFF-to-ON, and LORO scenarios, joined by LFP-derived Hjorth complexity and the temporal derivative of the beta/gamma ratio, both consistently pro-dystonic. Notably, the magnitude of LFP-derived Hjorth mobility was substantially larger in the OFF-to-ON LFP-only model (≈2.5) than in the Recording (≈1.4) or LORO Mixed (≈0.9), despite the OFF-to-ON scenario being the one in which the LFP classifier performed closest to chance level; this discrepancy between coefficient magnitude and predictive performance suggests that, in the absence of a second, complementary modality, the LFP-only model may over-weight this single feature. In the EEG-only model, central relative alpha power was invariably the top-ranked feature (pro-non-dystonic) in all three scenarios, with its first-order temporal derivative and central-temporal gamma-band coherence consistently pro-dystonic, and the EEG theta/beta ratio appearing among the top predictors throughout, mirroring the Combined-model results.

The fact that the Combined model's leading features are essentially the union of the leading LFP and EEG features, rather than a distinct set of cross-modal predictors, suggests that most of the discriminative power in the Combined model derives from each modality's own within-band dynamics rather than from EEG-LFP cross-coherence, with central-temporal gamma-band coherence standing out as the only cross-hemispheric coherence feature to consistently rank among the top predictors across all scenarios and both the Combined and EEG-only models.

Because gamma-band EEG activity is the frequency range most exposed to cranial muscle and movement contamination, and central-temporal gamma coherence ranked among the strongest EEG predictors, we performed a control analysis to test whether the classifiers depended on these features. The entire pipeline was repeated after removing all EEG gamma-band features, leaving all other features unchanged; full results are reported in the Supplementary Materials (Supplementary Section S1, Tables S1 and S2). In the mixed LORO scenario, the Combined model retained a balanced accuracy of 0.72 (vs. 0.74 with the full feature set) and detected 45/51 events (88.2% vs. 90.2%), with 76 false alarms (vs. 88). Its operating point became more conservative, with lower sensitivity (0.67 vs. 0.78) and higher specificity (0.78 vs. 0.70). The EEG-only model changed little (balanced accuracy 0.66 vs. 0.67), consistent with central alpha power being its leading predictor. Similar results were observed in the OFF-to-ON scenario (EEG-only 0.67 vs. 0.70; Combined 0.63 vs. 0.64), with event-level detection rates of 88.2% for both EEG and Combined models (Table S2). In the Recording scenario, the effect was larger (EEG-only 0.59 vs. 0.62; Combined 0.58 vs. 0.61), and the Combined model detected fewer events (44.1% vs. 50.0%). LFP-only results were identical to the main analysis by construction.

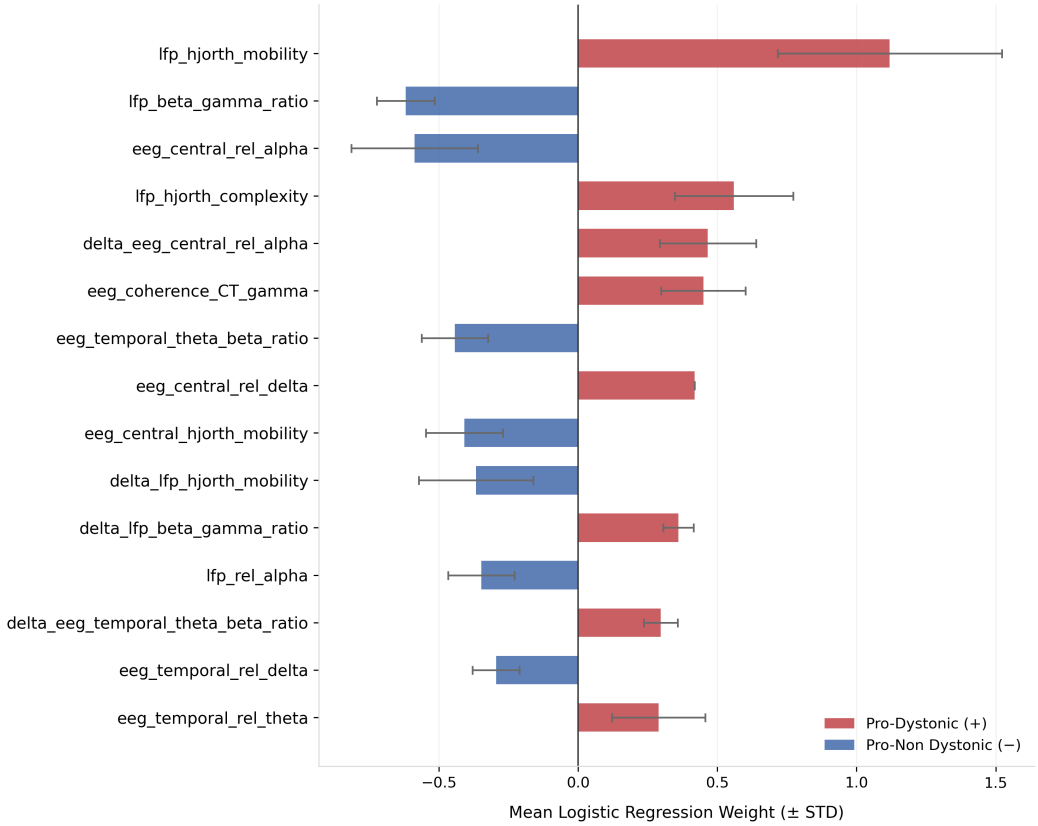


Figure 6. Top−15 feature importance for the Combined model in the mixed leave-one-recording-out (LORO) scenario, ranked by the absolute value of the mean logistic regression coefficient across folds. Red bars denote features positively associated with dystonic muscle activity; blue bars denote features associated with the non-dystonic condition. Error bars represent the between-fold standard deviation.

3.4. Generalization to Unseen Stimulation Intensities

As a further, more stringent test of generalization, we trained the Combined model on the same pool of recordings used in the mixed LORO scenario (recordings 1, 3, 5, and 6, the OFF condition and the 1.6 mA directional configuration) and evaluated it on recordings 4 and 7, which were held out entirely during training and correspond to stimulation intensities never encountered during model fitting (3.2 mA and 2.4 mA, respectively). Unlike the OFF-to-ON scenario, which probes generalization to a previously unseen stimulator state, this test probes generalization across current intensity itself.

The two held-out recordings showed a similar asymmetry to that reported for the OFF-to-ON scenario: recording 4 (3.2 mA) yielded a more balanced operating point (balanced accuracy 0.68, sensitivity 0.61, specificity 0.76 at $w = 5$), while recording 7 (2.4 mA) showed a more specificity-dominated profile (balanced accuracy 0.67, sensitivity 0.52, specificity 0.83 at $w = 5$), consistent with a general tendency toward under-detection of dystonic epochs when generalizing to unseen conditions.

Pooled across the two held-out recordings, the model achieved a balanced accuracy of 0.66 on raw predictions, rising to 0.69 after 5-epoch median smoothing and 0.70 with a 7-epoch window (Table 5, Figure 7), consistent with the pattern already observed in Section 3.1, whereby smoothing improves balanced accuracy primarily through gains in specificity ($0.75 \rightarrow 0.83$ from raw to $w = 7$) rather than sensitivity, which remained comparatively stable (0.57–0.58) across smoothing windows.

The logistic regression coefficients underlying this model were consistent with those already described for the LORO Combined model: LFP-derived Hjorth mobility and the LFP beta/gamma power ratio remained the strongest pro-dystonic and pro-non-dystonic

predictors, respectively, and central-temporal gamma-band coherence again ranked among the top pro-dystonic features.

Table 5. Epoch-level classification performance for the Combined model, trained on the LORO recording pool (OFF and 1.6 mA) and tested on previously unseen stimulation intensities (recordings 4 and 7; 3.2 mA and 2.4 mA), pooled across both recordings, by smoothing window.

Smoothing Window	Bal. Acc.	Sensitivity	Specificity	F1
1 (raw)	0.66	0.57	0.75	0.53
3	0.68	0.58	0.78	0.56
5	0.69	0.58	0.80	0.58
7	0.70	0.58	0.83	0.59

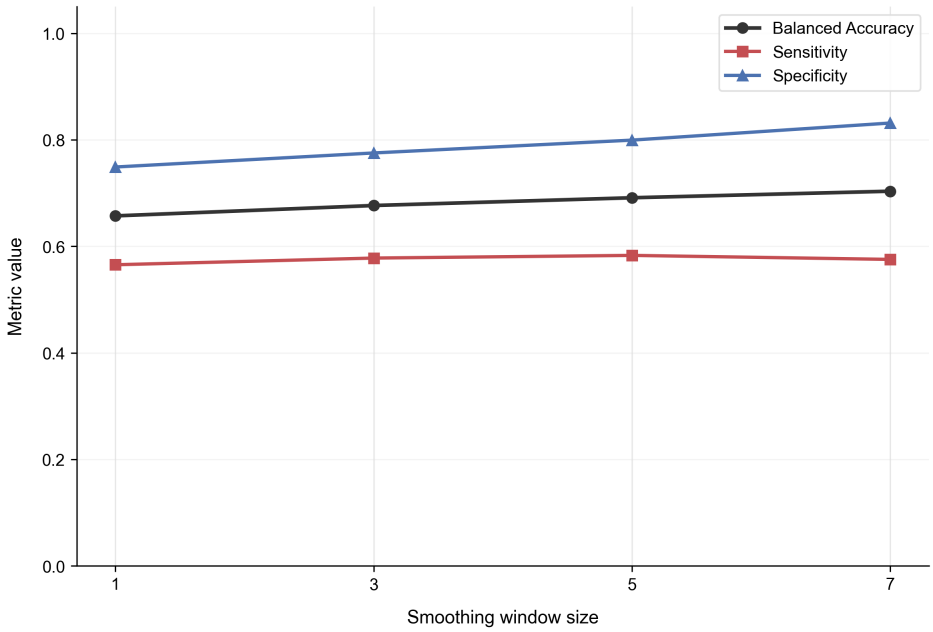


Figure 7. Balanced accuracy, sensitivity, and specificity as a function of the median-filter smoothing window, for the Combined model trained on the LORO recording pool and tested on previously unseen stimulation intensities (recordings 4 and 7).

4. Discussion

Programming DBS parameters in pediatric dystonia remains largely empirical, as no validated electrophysiological biomarker currently exists to guide their selection. Overall, these findings suggest that dystonic muscle activity can be distinguished from non-dystonic ones using synchronized EEG and LFP recordings in a pediatric patient with GPi-DBS; combining the two modalities improved balanced accuracy only in the mixed LORO scenario, and EEG alone performed equally well or better in the others.

LFP-derived classifiers exhibited a consistent bias across all three evaluation scenarios, with specificity systematically exceeding sensitivity, most notably in the Recording and LORO Mixed scenarios (0.75 vs. 0.44 and 0.79 vs. 0.47, respectively) and to a lesser extent in the OFF-to-ON scenario (0.61 vs. 0.51). This pattern suggests that the LFP-derived signature, dominated by Hjorth mobility and the beta/gamma power ratio, is a conservative predictor that reliably avoids false positives at the expense of under-detection of true dystonic muscle activity.

EEG-derived classifiers did not display an equally stable pattern. In the Recording and OFF-to-ON scenarios, EEG likewise showed higher specificity than sensitivity (0.70 vs. 0.54 and 0.75 vs. 0.64), although the gap was narrower than for LFP. In the mixed LORO

scenario, however, this relationship was reversed, with sensitivity reaching 0.80 against a specificity of 0.55. Notably, this is the scenario in which the Combined model achieved its best overall performance (balanced accuracy 0.74, sensitivity 0.78, specificity 0.70). This suggests that the benefit of the Combined model in the LORO scenario derives specifically from the convergence of a sensitivity-dominated EEG classifier with a specificity-dominated LFP classifier. This interpretation is further supported by the feature importance analysis, which showed that the leading predictors of the Combined model largely reproduced the union of the top EEG-only and LFP-only predictors, rather than introducing a distinct set of cross-modal terms.

A physiological basis for this complementarity lies in the different origin and spatial scale of the two signals. Scalp EEG is a spatially smoothed measure of the synchronous synaptic activity of aligned cortical populations, integrated over several square centimeters [17], whereas the pallidal LFP reflects local activity within a subcortical nucleus. Since dystonia involves a distributed basal ganglia–thalamo–cortical–cerebellar network, in which pallidal activity is coupled to sensorimotor, temporal and cerebellar cortex in distinct frequency bands [33], the two signals sample different nodes of the same network and can carry partly independent information. This is in line with our results: the EEG and LFP classifiers showed different error profiles, and the leading predictors of the Combined model were the union of the leading features of each modality rather than cross-modal coherence features. However, the advantage of the Combined model was nominal and limited to the mixed LORO scenario.

The mixed LORO scenario, which combined OFF and ON recordings and evaluated performance on entirely unseen recording sessions, was the validation scheme most representative of a short-term, intra-subject evaluation context, and it yielded the best overall performance for every modality, most notably for the Combined model (balanced accuracy 0.74, Table 2). When the Combined model was instead evaluated on recordings acquired at stimulation intensities entirely absent from training, performance was only partially preserved: balanced accuracy fell to 0.66 on raw predictions and recovered to 0.69–0.70 after smoothing (Table 5), driven almost entirely by specificity (0.75 → 0.83).

The effect of median-filter smoothing itself differed systematically by modality, improving balanced accuracy most consistently for EEG (+0.02 to +0.05 across scenarios), having a negligible or slightly negative effect for LFP, and producing a modest improvement for the Combined model. The event-level results in Table 4 offer a plausible explanation for this difference: in the LORO scenario, EEG-derived classifier generated 80 false alarms against 51 true events, more than twice the 39 false alarms produced when using only LFP, while achieving a shorter mean detection delay (0.54 vs. 1.57 epochs). A classifier that reacts quickly but produces more false positives, as seems to be the case with EEG, exhibits the kind of error profile that a temporal filter can effectively correct; a slower but more conservative classifier, as the LFP-based one appears to be, has relatively little noise to remove and therefore little room for improvement without also compromising its few correct detections.

Finally, our findings should be considered in the context of the biomarkers described in adult dystonia. The best established of these, pallidal theta–alpha (4–12 Hz) power in isolated and cervical dystonia [36], cannot be assumed to be valid in children. Pediatric dystonia is dominated by acquired forms, pallidal firing differs between aetiologies in children [37], and pediatric pallidal recordings have shown prominent beta activity alongside low-frequency components, both in status dystonicus [26] and during voluntary movement in dystonic cerebral palsy [27]. As we analyzed a single patient and did not test adult biomarkers prospectively, our results can neither confirm nor refute their validity in children, which will require larger pediatric cohorts.

4.1. Limitations

Several limitations should be taken into account when interpreting these results.

- This is an $N = 1$ single-patient study; thus, the identified candidate predictive features may reflect patient-specific characteristics rather than population-wide physiological patterns. The analyzed 1-s epochs represent repeated intra-individual observations rather than statistically independent biological replicates. Although the LORO validation successfully prevents temporal data leakage and demonstrates short-term, within-patient robustness across stimulation states, future studies must establish whether these models generalize across different patients and remain longitudinally stable across multiple recording sessions.
- This limitation is further compounded by class imbalance in some recordings: in recording 3 (right hemisphere), for instance, only 8 of 348 total epochs were dystonic, which limits how reliably fold-specific coefficients could be estimated for that session.
- Stimulation intensity and active contact configuration were not fully decoupled in this protocol, meaning that effects attributed here to stimulation intensity may be partly confounded with contact location.
- Dystonic labels were derived from EMG and clinical annotation, and should be interpreted as qEMG-derived dystonic muscle activity rather than as a direct measure of dystonia. Co-contraction is not specific to dystonia, sustained activity due to spasticity cannot be fully excluded, and dystonic postures without co-contraction in the recorded muscles may have been missed.
- The LFP stream was synchronized to the EEG/EMG recordings through start and end stimulation pulses. A residual lag between streams could affect EEG-LFP coupling features and the alignment between LFP epochs and EMG-derived labels, and cross-modal connectivity metrics should therefore be interpreted with caution (Supplementary Section S3). Packet loss in the LFP was limited (Supplementary Table S3) but cannot be excluded as a source of discontinuity artifacts.
- The recordings were acquired consecutively within a single session, separated by intervals of at least 20 min. This does not constitute a clinical washout: residual effects of the preceding setting may have been present, and the full clinical effect of a new setting may not have developed within a 5-min block. The OFF and ON labels therefore indicate the programmed stimulation state rather than a steady clinical state. Leave-one-recording-out validation prevents temporal leakage between training and test data, but it cannot separate stimulation effects from order effects. The ability of the classifier to discriminate dystonic epochs across blocks suggests that the identified features are not specific to a single recording, but replication with randomized or counterbalanced stimulation order and adequate washout is needed.

These limitations do not undermine the main conclusion that a multimodal EEG-LFP decoder is capable of distinguishing dystonic activity from non-dystonic activity. However, they imply that the candidate features identified here, LFP Hjorth mobility and the beta/gamma power ratio, central EEG alpha power and its temporal derivative, as well as central-temporal gamma coherence, should be considered promising candidates for further investigation in a larger cohort.

4.2. Future Directions

Future work should first replicate this pipeline in a larger cohort of pediatric patients undergoing GPi-DBS, to determine whether the features identified here generalize across patients. Explicitly incorporating stimulation amplitude and contact configuration as continuous model inputs, rather than as scenario-defining variables, may help reduce the generalization gap observed here. Longer recordings would then show whether the com-

plementarity between EEG and LFP remains stable over time, and would provide the larger samples needed to separate stimulation-state effects from the limits of small training sets. Finally, to test whether decoded dystonic activity can support DBS parameter selection, either offline, by comparing stimulation settings within a programming session, or online, as a control signal for adaptive DBS, studies with randomized stimulation order, adequate washout and clinical outcome measures will be needed.

5. Conclusions

In a single pediatric patient with generalized secondary dystonia and bilateral GPi-DBS, dystonic muscle activity could be discriminated from non-dystonic ones using synchronized EEG and LFP recordings, with all modalities performing significantly above a label-permutation null; the Combined EEG-LFP model achieved the highest balanced accuracy in the mixed leave-one-recording-out scenario, with its most favorable balance of sensitivity and specificity (0.78 and 0.70, respectively), although its advantage over EEG alone was nominal and did not survive multiple-comparison correction. This performance did not stem from the identification of novel cross-modal signals, but from the complementary error profiles of the two modalities: LFP-derived classifiers were consistently specificity-dominated across all scenarios, whereas EEG-derived classifiers became markedly sensitivity-dominated specifically in the mixed LORO scenario, the condition under which the Combined model performed best. A small set of features (LFP Hjorth mobility and the beta/gamma power ratio, EEG central alpha power and its temporal derivative, and central-temporal gamma coherence) recurred consistently across validation schemes and single-modality models. Generalization to unseen stimulation states and unseen stimulation intensities was only partially preserved, particularly for LFP-derived features. These results should be interpreted as methodological proof-of-concept evidence, warranting further investigation of multimodal electrophysiological correlates of dystonic activity in pediatric patients, but requiring replication in larger cohorts.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/bioengineering13111313/s1, Figure S1: Absolute LFP power spectral density in dystonic (red) and non-dystonic (blue) epochs, pooled across recording/hemisphere series. (A) Stimulator OFF; (B) stimulator ON (1.6 mA). Both axes are logarithmic. Shaded bands: 95% block-bootstrap confidence intervals. Grey area: frequencies outside the specparam fit range (3–40 Hz); Figure S2: Specparam fits of the pooled LFP spectra in the stimulator-OFF condition (fit range 3–40 Hz, fixed aperiodic mode). Left: dystonic epochs; right: non-dystonic epochs. Black: original spectrum; red: full model; dashed blue: aperiodic component; Figure S3: Specparam fits of the pooled LFP spectra in the stimulator-ON condition (1.6 mA; fit range 3–40 Hz, fixed aperiodic mode). Left: dystonic epochs; right: non-dystonic epochs. Black: original spectrum; red: full model; dashed blue: aperiodic component; Figure S4: Power spectral density before (grey) and after (blue) preprocessing during the stimulator-ON condition (Recording 3). Top row (left hemisphere): (A) EEG central channel C3, (B) EEG temporal channel T3, and (C) left GPi LFP. Bottom row (right hemisphere): (D) EEG central channel C4, (E) EEG temporal channel T4, and (F) right GPi LFP. Insets display a magnification of the 20–45 Hz frequency range. Vertical dotted orange lines denote the empirical notch filters applied to EEG channels (29.0 and 37.8 Hz), and the shaded blue area indicates the analysis bandpass (0.5–45 Hz). Absolute PSD is expressed in $\mu V^2/Hz$ on a logarithmic scale; Table S1: Epoch-level classification performance after excluding all EEG gamma-band features (Logistic Regression; raw → smoothed, $w = 5$). Metrics are pooled across folds. LFP-only results are identical to those of the main analysis because LFP features were not modified.; Table S2: Event-level detection performance after excluding all EEG gamma-band features (smoothed predictions); Table S3: LFP telemetry: valid packets, lost packets and interruptions for each recording (percentages relative to the total number of packets). Recording 2 was not used in the analyses; Table S4: Epoch-level performance by hemisphere (L, left; R, right), smoothed predictions ($w = 5$). Metrics are computed

on the out-of-fold predictions of each scenario pooled across recordings. Dystonic/total: number of dystonic epochs over the total number of epochs in the subset; Table S5: Epoch-level performance of the Combined model for each recording/hemisphere series (L, left; R, right), smoothed predictions ($w = 5$). Each recording is shown in the scenario in which it was a test set. Sensitivity and balanced accuracy for Recording 3, right hemisphere, rely on 8 dystonic epochs only.

Author Contributions: Conceptualization, E.B., E.F., C.R., L.M. and M.L. (Matteo Lenge); methodology, E.B. and C.R.; software, E.B. and C.R.; validation, E.B., C.R., E.F., A.D.R., S.P., A.N., S.C., M.L. (Marta Lazzeri), F.M., M.L. (Matteo Lenge), F.G. and L.M.; formal analysis, E.B.; investigation, E.B., C.R., E.F., A.D.R., S.P., A.N., S.C., M.L. (Marta Lazzeri), F.M., M.L. (Matteo Lenge), F.G. and L.M.; resources, E.B., C.R., E.F., A.D.R., S.P., A.N., S.C., M.L. (Marta Lazzeri), F.M., M.L. (Matteo Lenge), F.G. and L.M.; data curation, E.B., C.R., E.F., A.D.R., S.P., A.N., S.C., M.L. (Marta Lazzeri), F.M., M.L. (Matteo Lenge), F.G. and L.M.; writing—original draft preparation, E.B.; writing—review and editing, E.B., C.R., E.F., A.D.R., S.P., A.N., S.C., M.L. (Marta Lazzeri), F.M., M.L. (Matteo Lenge), F.G. and L.M.; visualization, E.B.; supervision, F.G., M.L. (Matteo Lenge) and L.M.; project administration, F.G., M.L. (Matteo Lenge) and L.M.; funding acquisition, F.G., M.L. (Matteo Lenge) and L.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the National Plan for Complementary Investments to the PNRR Hub Life Science-Digital Health (PNC-E3-2022-23683267, DHEAL COM). This publication reflects only the authors' view and the Italian Ministry of Health is not responsible for any use that may be made of the information it contains.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Ethics Committee of Pediatric Ethics Committee of the Tuscany Region, Italy (protocol code 33/2025 and 13/02/2025 of approval).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Data are available on reasonable request from the corresponding author due to privacy and ethical restrictions.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

CI	Confidence Interval
DBS	Deep Brain Stimulation
EEG	Electroencephalography
EMG	Electromyography
GPe	Globus Pallidus externus
GPi	Globus Pallidus internus
ICA	Independent Component Analysis
IIR	Infinite Impulse Response
LFP	Local Field Potential
LORO	Leave-One-Recording-Out
PSD	Power Spectral Density
qEMG	quantitative Electromyography

References

1. Soliman, Y.; Figliomeni, A.; Lee, H.; Thorburn, M.; Ghosh, S.; Ramakonar, H.; Yau, W.Y. Deep brain stimulation for the treatment of paediatric dystonia due to neuroanatomical abnormalities: A systematic review of efficacy and future perspectives. *BMJ Neurol. Open* **2026**, *8*, e001543. [CrossRef] [PubMed]
2. Albanese, A.; Bhatia, K.P.; Fung, V.S.; Hallett, M.; Jankovic, J.; Klein, C.; Krauss, J.K.; Lang, A.E.; Mink, J.W.; Pandey, S.; et al. Definition and classification of dystonia. *Mov. Disord.* **2025**, *40*, 1248–1259. [CrossRef] [PubMed]
3. Albanese, A.; Bhatia, K.; Bressman, S.B.; DeLong, M.R.; Fahn, S.; Fung, V.S.; Hallett, M.; Jankovic, J.; Jinnah, H.A.; Klein, C.; et al. Phenomenology and classification of dystonia: A consensus update. *Mov. Disord.* **2013**, *28*, 863–873. [CrossRef] [PubMed]

4. Sanger, T.D.; Chen, D.; Fehlings, D.L.; Hallett, M.; Lang, A.E.; Mink, J.W.; Singer, H.S.; Alter, K.; Ben-Pazi, H.; Butler, E.E.; et al. Definition and classification of hyperkinetic movements in childhood. *Mov. Disord.* **2010**, *25*, 1538–1549. [CrossRef] [PubMed]
5. Fahn, S. Classification of movement disorders. *Mov. Disord.* **2011**, *26*, 947–957. [CrossRef] [PubMed]
6. Neychev, V.K.; Gross, R.E.; Lehéricy, S.; Hess, E.J.; Jinnah, H. The functional neuroanatomy of dystonia. *Neurobiol. Dis.* **2011**, *42*, 185–201. [CrossRef] [PubMed]
7. Piña-Fuentes, D.; Beudel, M.; Little, S.; van Zijl, J.; Elting, J.W.; Oterdoom, D.M.; van Egmond, M.E.; van Dijk, J.M.C.; Tijssen, M.A. Toward adaptive deep brain stimulation for dystonia. *Neurosurg. Focus* **2018**, *45*, E3. [CrossRef] [PubMed]
8. Mink, J.W. Special concerns in defining, studying, and treating dystonia in children. *Mov. Disord.* **2013**, *28*, 921–925. [CrossRef] [PubMed]
9. Lin, J.P.; Lumsden, D.E.; Gimeno, H.; Kaminska, M. The impact and prognosis for dystonia in childhood including dystonic cerebral palsy: A clinical and demographic tertiary cohort study. *J. Neurol. Neurosurg. Psychiatry* **2014**, *85*, 1239–1244. [CrossRef] [PubMed]
10. Elkaim, L.M.; Alotaibi, N.M.; Sigal, A.; Alotaibi, H.M.; Lipsman, N.; Kalia, S.K.; Fehlings, D.L.; Lozano, A.M.; Ibrahim, G.M.; Collaboration, N.A.P.D.; et al. Deep brain stimulation for pediatric dystonia: A meta-analysis with individual participant data. *Dev. Med. Child Neurol.* **2019**, *61*, 49–56. [CrossRef] [PubMed]
11. Marks, W.; Bailey, L.; Reed, M.; Pomykal, A.; Mercer, M.; Macomber, D.; Acosta, F., Jr.; Honeycutt, J. Pallidal stimulation in children: Comparison between cerebral palsy and DYT1 dystonia. *J. Child Neurol.* **2013**, *28*, 840–848. [PubMed]
12. McQueen, S.A.; Gorodetsky, C.; Ibrahim, G.M. Deep brain stimulation for neurological and neurodevelopmental disorders in children: Current applications and future perspectives. *Dev. Med. Child Neurol.* **2025**, *67*, 1536–1542. [CrossRef] [PubMed]
13. Magown, P.; Andrade, R.A.; Soroceanu, A.; Kiss, Z.H. Deep brain stimulation parameters for dystonia: A systematic review. *Park. Relat. Disord.* **2018**, *54*, 9–16. [CrossRef] [PubMed]
14. Johnson, K.A.; Coutinho, P.B.; Sarmento, F.P.; Vo, E.; Wong, J.K.; Hilliard, J.D.; Foote, K.D.; de Hemptinne, C. Pallidal evoked resonant neural activity as a biomarker for deep brain stimulation in dystonia. *medRxiv* **2025**. [CrossRef]
15. Jung, Y.; Mithani, K.; Suresh, H.; Warsi, N.; Harmsen, I.E.; Breitbart, S.; Gorodetsky, C.; Fasano, A.; Fallah, A.; Hadjinicolaou, A.; et al. Deep brain stimulation in pediatric populations: A scoping review of the clinical trial landscape. *Stereotact. Funct. Neurosurg.* **2025**, *103*, 132–144. [CrossRef] [PubMed]
16. Cohen, M.X. Where does EEG come from and what does it mean? *Trends Neurosci.* **2017**, *40*, 208–218. [CrossRef] [PubMed]
17. Buzsáki, G.; Anastassiou, C.A.; Koch, C. The origin of extracellular fields and currents—EEG, ECoG, LFP and spikes. *Nat. Rev. Neurosci.* **2012**, *13*, 407–420. [CrossRef] [PubMed]
18. Ebden, M.; Elkaim, L.M.; Breitbart, S.; Yan, H.; Warsi, N.; Huynh, M.; Mithani, K.; Gouveia, F.V.; Fasano, A.; Ibrahim, G.M.; et al. Chronic Pallidal local field potentials are associated with dystonic symptoms in children. *Neuromodul. Technol. Neural Interface* **2024**, *27*, 551–556. [CrossRef] [PubMed]
19. Ledingham, D.; Baker, M.; Pavese, N. Local field potentials: Therapeutic implications for DBS in dystonia including adaptive DBS for dystonia. *Deep Brain Stimul.* **2024**, *5*, 4–19. [CrossRef]
20. Kasiri, M.; Asadi, M.; Nataraj, J.; Abe, S.; Sanger, T.D. Optimized dbs in pediatric dystonia restores balance in transmission of signals within pallidothalamic network by modulating neural oscillations in deep brain regions. *medRxiv* **2025**. [CrossRef]
21. Vicente, R.; Wibral, M.; Lindner, M.; Pipa, G. Transfer entropy—A model-free measure of effective connectivity for the neurosciences. *J. Comput. Neurosci.* **2011**, *30*, 45–67. [CrossRef] [PubMed]
22. Cernera, S.; Shcherbakova, M.; Hammer, L.H.; Friedrich, M.; Peach, R.; Ip, C.W.; Bledsoe, I.; Little, S.; Starr, P.A. Oscillatory dynamics in isolated dystonia: Five hundred hours of chronic invasive multisite motor network recordings. *J. Neurophysiol.* **2025**, *134*, 677–690. [CrossRef] [PubMed]
23. Bahners, B.H.; Lofredi, R.; Voss, H.; de Almeida Marcelino, A.L.; Goede, L.L.; Feldmann, L.K.; Schnitzler, A.; Sander, T.H.; Florin, E.; Kühn, A.A. Spatial signature of low-frequency network changes accounts for pallidal stimulation outcome in cervical dystonia. *EBioMedicine* **2026**, *124*, 106140. [CrossRef] [PubMed]
24. Geng, X.; Quan, Z.; Zhang, R.; Zhu, G.; Nie, Y.; Wang, S.; Rolls, E.; Zhang, J.; Hu, L. Subthalamic and pallidal oscillations and their couplings reflect dystonia severity and improvements by deep brain stimulation. *Neurobiol. Dis.* **2024**, *199*, 106581. [CrossRef] [PubMed]
25. Horan, J.; Donlon, E.; Gill, A.; Moran, R.; Fearon, C.; Walsh, R.A.; Moran, C. Acute and Chronic Local Field Potential Recordings in Dystonia—A Systematic Review. *Mov. Disord. Clin. Pract.* **2026**, *early view*. [CrossRef] [PubMed]
26. Balachandar, A.; Vogt, L.M.; Mithani, K.; Coleman, S.C.; Ebden, M.; Leblanc-Miller, A.; Breitbart, S.; Fasano, A.; Gorodetsky, C.; Ibrahim, G.M. Status dystonicus is a distinct state characterized by pallidal beta-band activity. *Nat. Commun.* **2025**, *16*, 9352. [CrossRef] [PubMed]
27. Hernandez-Martin, E.; Kasiri, M.; Abe, S.; MacLean, J.; Olaya, J.; Liker, M.; Chu, J.; Sanger, T.D. Globus pallidus internus activity increases during voluntary movement in children with dystonia. *iScience* **2023**, *26*, 107066. [CrossRef] [PubMed]

28. Neumann, W.J.; Gilron, R.; Little, S.; Tinkhauser, G. Adaptive deep brain stimulation: From experimental evidence toward practical implementation. *Mov. Disord.* **2023**, *38*, 937–948. [CrossRef] [PubMed]
29. Oliveira, A.M.; Coelho, L.; Carvalho, E.; Ferreira-Pinto, M.J.; Vaz, R.; Aguiar, P. Machine learning for adaptive deep brain stimulation in Parkinson's disease: Closing the loop. *J. Neurol.* **2023**, *270*, 5313–5326. [CrossRef] [PubMed]
30. Horn, A.; Kühn, A.A. Lead-DBS: A toolbox for deep brain stimulation electrode localizations and visualizations. *Neuroimage* **2015**, *107*, 127–135. [CrossRef] [PubMed]
31. Barbi, E.; Di Rita, A.; Mesin, L.; Rossi, C.; Noris, A.; Guerrini, R.; Fino, E.; Ciabattini, S.; Giordano, F.; Peraio, S.; et al. Surface and Deep Signal Correlation in a Pediatric Dystonia Case with DBS. In *Proceedings of the 2025 IEEE International Conference on Metrology for eXtended Reality, Artificial Intelligence and Neural Engineering (MetroXRAINE)*; IEEE: New York, NY, USA, 2025; pp. 966–971.
32. Homan, R.W.; Herman, J.; Purdy, P. Cerebral location of international 10–20 system electrode placement. *Electroencephalogr. Clin. Neurophysiol.* **1987**, *66*, 376–382. [CrossRef] [PubMed]
33. Neumann, W.J.; Jha, A.; Bock, A.; Huebl, J.; Horn, A.; Schneider, G.H.; Sander, T.H.; Litvak, V.; Kühn, A.A. Cortico-pallidal oscillatory connectivity in patients with dystonia. *Brain* **2015**, *138*, 1894–1906. [CrossRef] [PubMed]
34. Welch, P. The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. *IEEE Trans. Audio Electroacoust.* **1967**, *15*, 70–73. [CrossRef]
35. Hjorth, B. EEG analysis based on time domain properties. *Electroencephalogr. Clin. Neurophysiol.* **1970**, *29*, 306–310. [CrossRef] [PubMed]
36. Neumann, W.J.; Horn, A.; Ewert, S.; Huebl, J.; Brücke, C.; Slentz, C.; Schneider, G.H.; Kühn, A.A. A localized pallidal physiomarker in cervical dystonia. *Ann. Neurol.* **2017**, *82*, 912–924. [CrossRef] [PubMed]
37. McClelland, V.; Valentin, A.; Rey, H.G.; Lumsden, D.; Elze, M.; Selway, R.; Alarcon, G.; Lin, J. Differences in globus pallidus neuronal firing rates and patterns relate to different disease biology in children with dystonia. *J. Neurol. Neurosurg. Psychiatry* **2016**, *87*, 958–967. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.