

# Mapping the electrophysiological structure of dystonic Globus Pallidus pars interna through intraoperative microelectrode recordings

Ahmet Kaymak<sup>1,2</sup>, Matteo Vissani<sup>3</sup>, Sara Rinaldo<sup>4</sup>, Roberto Eleopra<sup>4</sup>, Luigi Romito<sup>4</sup>,  
Alberto Mazzoni<sup>1,2</sup>

<sup>1</sup>The Biorobotics Institute, Scuola Superiore Sant'Anna, Pisa 56025, Italy

<sup>2</sup>Department of Excellence for Robotics and AI, Scuola Superiore Sant'Anna, Pisa 56025, Italy

<sup>3</sup>Department of Neurosurgery, Massachusetts General Hospital, Boston 02115, United States of America

<sup>4</sup>Movement Disorders Department, Fondazione IRCCS Istituto Neurologico Carlo Besta, 20133 Milan, Italy

Corresponding Author

Dr. Alberto Mazzoni, The Biorobotics Institute of Scuola Superiore Sant'Anna, Viale Rinaldo Piaggio  
34, Pontedera, Pisa, Italy

E-mail: [alberto.mazzoni@santannapisa.it](mailto:alberto.mazzoni@santannapisa.it)

## Abstract

*Objective.* The Globus Pallidus pars interna (GPi) is one of the main targets for Deep Brain Stimulation (DBS) therapies for dystonia and other movement disorders. Still, a complete picture of the spiking dynamics of the nucleus is far from being achieved. Microelectrode recordings (MER) provide a unique brain window opportunity to shed light on GPi organization, which might support intraoperative DBS target localization, as previously done for the Subthalamic nucleus (STN). *Approach.* Here we propose a novel procedure to analyze explorative MERs from DBS implants in dystonic patients. The procedure identifies the neural activity markers discriminating neurons in the GPi from those in the neighbouring structures, as well as the markers discriminating neurons located in different regions within the GPi. *Main results.* The identification of the borders of the GPi based on neural markers was a difficult task, due to internal inhomogeneities in GPi firing dynamics. However, the procedure was able to exploit these inhomogeneities to characterize the internal electrophysiological structure of the GPi. In particular, we found a reliable dorsolateral gradient in firing activity and regularity. *Significance.* Overall, we characterized the spatial distribution of neural activity markers in the dystonic GPi, paving the way for the use of these markers for DBS target localization. The procedure we developed to achieve this result could be easily extended to MER performed for other disorders and in other areas.

Keywords: Dystonia, Globus Pallidus Pars Interna (GPi), Microelectrode Recordings, Deep Brain Stimulation (DBS), Stimulation Target Localization, Movement Disorders

## Introduction

Deep brain stimulation (DBS) of the internal segment of the globus pallidus (GPi) has been established as a highly therapeutic procedure for dystonia. However, few research studies have been carried

[1–3]. The inappropriate location of the active contacts explains most of the therapeutical failures in GPi DBS for dystonia, which increases the rate of revisions and remissions of previously implanted DBS leads [4,5]. Maximizing the overlap between the GPi optimal location and the 3D spread

of the current delivered by the active contact of the DBS lead is critical to achieving clinically significant outcomes [6,7].

Although advances in imaging protocols have significantly improved preoperative GPi targeting, methodological limitations – such as low spatial resolution and possible brain shifts during the surgery – owe the utilization of up-to-five microelectrode recordings (MER) to intraoperatively validate the target location [8–10]. However, MERs can be time-consuming, which limits their use if an expert electrophysiologist is not available in the operating room. Scaling up the number of microelectrodes used during electrophysiological mapping allows for gaining more information, but also increases the risk of complications [11].

Multiple human intraoperative centres showed that MERs yield accurate electrophysiological spatial demarcation of target borders and identification of neural patterns that highlight optimal stimulation settings as well as discriminate clinical states [12–20]. Most of these studies focused on the Subthalamic nucleus (STN) *per se*, which represents another important DBS target for dystonia and other movement disorders, such as Parkinson’s Disease (PD) [21,22].

The peculiar characteristics of the transition from the external segment of the globus pallidus (GPe) to GPi, such as the sparseness of the neurons along the pallidum, the absence of a marked reduction in neuronal activity, and the high discharge rate variability, challenged the electrophysiological determination of GPi borders [27,28]. Occasionally, the identification of increasing background activity and *border cells*, characterized by low firing rates (2 – 20 Hz), may help identify the GPe-GPi transition.

Multiple MERs studies compared the activity of GPi neurons either across different diseases or before, during, and after functional microstimulation [24–31], neglecting or restricting the topographical analysis along the pre-planned trajectory axis. These studies identified *high-frequency cells* whose discharge rate (up to 150 Hz) may feature movement disorders. It is assumed that PD GPi cells are characterized by a higher discharge rate than dystonia GPi cells. Low-frequency oscillatory activity (2 – 10 Hz) and bursting patterns are also observed in GPi across different diseases. All these findings may be confounded by different anaesthetic protocols and the heterogeneity of neural features (both discharge rate and discharge pattern) in different Dystonic phenotypes. Strikingly, a topographic map of the spiking dynamics of GPi neurons has not been elucidated yet. Clinically-informed decisions may leverage these topographic maps to define relationships between neural activity and optimal stimulation sites.

Interestingly, complementary approaches using postoperative local field potential (LFP) recordings showed excessive entrainment of spatially localized theta oscillations in the Dystonic GPi, indexing the optimal stimulation site [32].

Here, our aim is to provide an MER-based machine-learning pipeline that retrospectively demarcates electrophysiologically-defined GPi borders and builds a topographic map of the internal GPi spiking dynamics in Dystonic patients. This work is in line with the growing interest in electrophysiological GPi physiome markers and might lay the bedrock of the implementation of future computer-aid systems that support and hasten clinical decisions during the placement of the DBS lead in GPi DBS dystonic patients [17,32,33]. The development of algorithms that interpret short segments of MERs during DBS surgery is crucial to maximally exploit the limited time in operating room settings.

## Methods

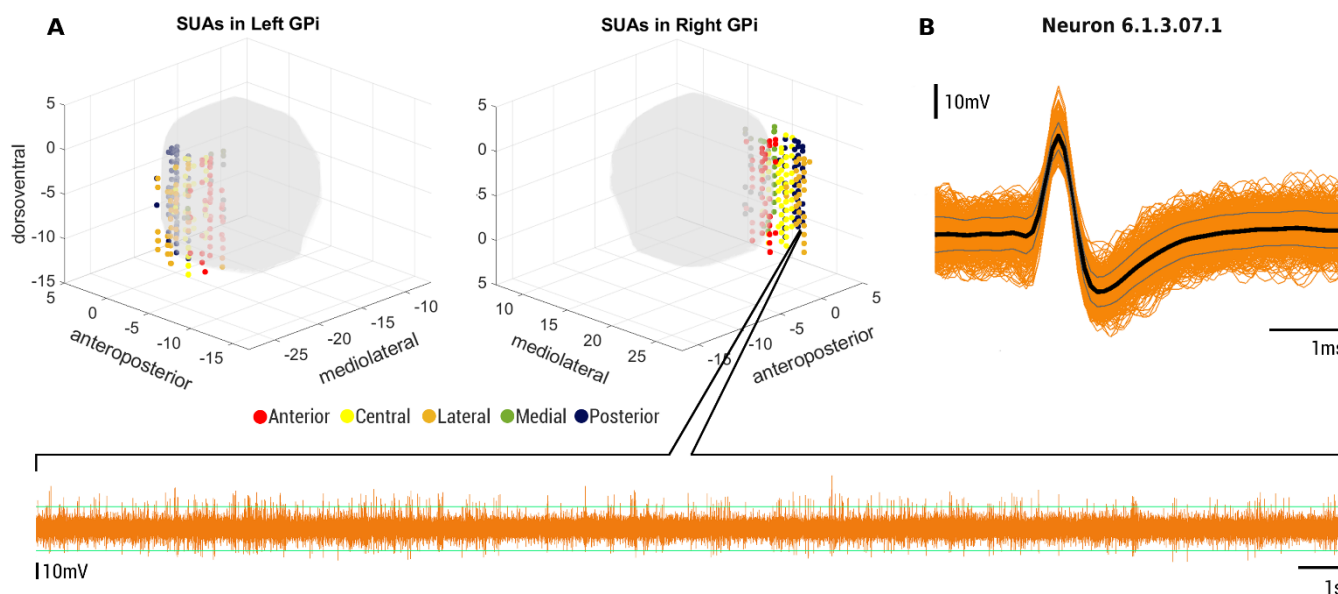
### 2.1 Subjects

The research activity was conducted with seven male and three female patients who are diagnosed with primary dystonia. All the patients who took part in this study had severe dystonia, which significantly limited their ability to carry out their daily tasks. Patients were addressed to surgery relying on these inclusion criteria: diagnosis of primary generalized or multisegmental dystonia, normal behavioural, psychiatric, and cognitive profile, normal brain MRI and lack of response to medical treatment including anticholinergics, benzodiazepines, neuroleptics, baclofen and/or botulinum toxin injections. Patients underwent bilateral implantation of quadripolar electrodes, after giving their written informed consent to the procedure.

### 2.2 Surgery

The patient underwent frameless surgery. Preoperative images (MRI and CT with the fiducials on the skull) and planning were acquired the day before surgery. Non-stereotactic MRI brain scans (T1 with gadolinium and T2; slice thickness: 1.5 mm; without gap or overlap) were performed and, after the MRI acquisition, metal screw fiducials were fixed to the skull of each patient under local anaesthesia. Thereafter, a volumetric non-contrast scan (axial, slice thickness from 0.75 to 1 mm) with the fiducials in place was performed. For all the procedures, CT data were matched with the previous MRI datasets in the Medtronic Stealth Station (Framelink®; Medtronic, Minneapolis, MN).

To define the anatomic target, we used fixed distances from the midpoint of the AC-PC line on T1 MRI slices combined with the direct visualization of the nucleus on T2 MRI images, to identify the brain target (GPi). In our study, we did not have medical imaging that belongs to GPi DBS surgeries of patients to inspect the boundaries of GPi. Hence there is no certainty in terms of relative positions of MER trajectories to GPi for each subject. To overcome this issue, the DISTAL atlas[34] which is a three-dimensional



**Figure 1. Isolated single unit representations on GPI in MNI reference frame.** (A) Locations of the single-unit activities are displayed in two dystonic GPI based on their MNI coordinate definitions. Color-coded perpendicular lines along the dorsoventral axis of GPI represent electrode trajectories (anterior, central, lateral, medial, and posterior) GPI is depicted as defined by the DISTAL atlas. (B) In addition, a raw MER recording and its single-unit activity from -7mm (lateral trajectory) in the right hemisphere of patient 6 are presented.

subcortical brain atlas that also contains the anatomical position of GPI in Montreal Neurological Institute (MNI) space. Target coordinates derived from the stereotactic atlases registered to the mid-commissural point were manually entered into the planning program screen and identified onto the corresponding image slice. The atlas coordinates for the GPI nucleus were 21 mm lateral, 2 mm anterior, and 4 mm deep to the mid-AC-PC point. The anatomic coordinates of the preoperative target were finally defined by slightly adjusting the atlas coordinates and carefully evaluating details about the anatomic structures surrounding the estimated target, appreciable as a result of CT fusion with MRI. GPI definition was used later to determine whether sorted neurons are situated inside or outside of the GPI.

The patient was placed on the operating table, and the surgeon did the first navigation and marked the frontal entry points on the skin using a “passive frame”: referring to the fiducials screwed in the skull and assessing the correct position of the camera in the operating room, navigation on Stealth Station showed “geometry error”: geometry errors of the Nexprobe and the reference arc were less than 0.5mm in all the patients of this sample.

A drill hole of 14 mm in diameter was created 2.5 to 3.5 cm off the midline anterior to the coronal suture, and it was closed by a cap of fibrin sealant to avoid deliquoration. The NexFrame System® (Medtronic, Minneapolis, MN) was used for each procedure performed.

In the majority of the patients, surgery was performed while the patient was awake but, in those patients with severe dystonic fixed postures or those unable to tolerate an awake

surgical procedure we performed surgery under general anaesthesia with a ketamine-based anaesthetic protocol. This anaesthetic protocol did not alter the electrophysiologic basal ganglia activity, allowing MER to be achievable and reliable during DBS surgery[35–37].

The choice of final electrode position was based on multi-electrode recording qualitative evaluation by an expert neurophysiologist, clinical efficacy during macrostimulation, and the absence of significant macrostimulation-induced collateral side effects. MER and macrostimulation were performed using the Leadpoint System® (Medtronic, Minneapolis, MN). Finally, surgery was completed with the implantation of the subcutaneous extensions and the IPG in the sub-clavicle area.

#### 2.4 Electrode Localization

The day after the surgical procedure, the patient underwent postoperative volumetric non-contrast scans (axial, slice thickness 0.75-1 mm) CT scans, to evaluate the final lead position and assess postoperative pneumocephalus and rule out asymptomatic haemorrhage. The postoperative CT scan was merged with the preoperative CT and MRI images using Medtronic Stealth Station (Framelink®; Medtronic, Minneapolis, MN) to blend the AC-PC coordinates and angles of the trajectory of the final DBS leads. Each final lead location was identified as the centre of the beam-hardening artefact representing the deepest of the electrodes. This mapping of MER electrode trajectories to MNI space was a rigid transformation.

## 2.5 Electrophysiological recordings

Electrophysiological explorative microelectrode recordings were simultaneously recorded along five different trajectories: anterior, lateral, medial, central, and posterior trajectories. The recordings were collected bilaterally and simultaneously from the right and left hemispheres with varying depths from -10 mm to +3.5 mm (0 mm is the planned target point) by using FHC semi-micro-electrode (impedance=1 MOhm (FC2002, Medtronic R, Minneapolis, MN)). During the recording procedure, an analogue 500 Hz high pass filter was applied. MERs were sampled at 24000 Hz for 10 seconds. In both hemispheres, the MER trajectories are parallel to the GPi's dorsoventral axis, with the central trajectory crossing the DBS target (see Supplementary Material Section 2.1).

We examined 1640 raw microelectrode recordings of 10 patients (min. 99, max. 246,  $164 \pm 50.19$  recordings per subject) (see Supplementary Material Section 2). Putative single unit activities (SUAs) are successfully isolated from raw MERs (Figure 1A). SUA represents the section of the raw data that belongs to the spiking of a single neuron. The [-0.5 +2.5] ms time interval around the peak values of all spikes extracted from raw data and stored as the neural activity of a single neuron[38].

## 2.6 Computational pipeline

To analyse MERs, we implemented a computational pipeline where each module is specialized for a particular task. In total, we implemented four modules: (1) Offline Spike Sorting Module, (2) Candidate Biomarker Generation Module, (3) Mutual Information Analysis Module, and (4) Machine Learning Module (see Figure 2).

### 2.6.1 Spikes detection and sorting

In the offline spike sorting module (Figure 2, top, left box), spike detection and sorting operations were conducted (see Supplementary Material Section 3.2). For spike identification and sorting, we used Wave Clus MATLAB Toolbox[39]. Double-thresholding (4 standard deviations from the baseline signal) was used to designate spike peaks, and 64-wavelet coefficients were computed for each potential spike. Then, using a superparamagnetic clustering method, the spikes are clustered to specific neurons based on their wavelet coefficients. We decomposed raw recordings into three main components: single-unit activity (SUA), multi-unit activity (MUA), and background activity (BUA)[38](see Supplementary Material Section 3.3)[14,15]. Each of these components carries different types of information about surrounding neural activity[38]. Noteworthy, we applied additional stability criteria that helped us prevent the inclusion of candidate neurons that do

not reliably resemble SUA signatures for further analysis (see Supplementary Material Section 3.4).

### 2.6.2 Reference frames

We carried out all of our analyses in two different reference frames: 1D relative depth and 3D MNI reference frames. The definition of two distinct reference frames allowed us to utilize our localization tasks both for clinical aspects and for understanding the electrophysiological structures of GPi. The 1D reference frame denotes the direction that is primarily parallel to GPi's dorsoventral axis. During the GPi DBS surgeries, neurosurgeon teams penetrate the electrodes perpendicularly to GPi. Localization and real-time feedback in the 1D reference frame offer clinical utility from this standpoint.

Although the 1D reference frame has clinical relevancy, it falls short of accurately reflecting neural dynamics in terms of the actual anatomical structure of GPi. To address this problem, we used the MNI coordinates of SUAs to characterize the pathological dynamics of dystonic GPi and tried to localize neurons in the 3D space.

### 2.6.3 Candidate neural biomarkers

Even if recent advances have been made in GPi bordering[40,41] and functional segregation of GPi [37], these approaches are not adequate to provide real-time feedback to neurosurgeon teams during DBS procedures for target localization. We used the definition of candidate neural biomarkers to locate and characterize neural behaviour inside different areas of GPi and to distinguish GPi borders for patients with primary dystonia. Operatively, we categorized neural dynamics into five different groups: (a) firing rate and regularity, (b) neural oscillations, (c) neural bursts, (d) spectral coherence between SUA and BUA, and the power spectrum of single-unit activity in different frequency bands. Then in each group, we defined multiple candidate neural biomarkers that represent a particular aspect of the related neural phenomena (see Supplementary Material Section 4).

### 2.6.4 Power Spectrum Estimation

Firstly, we extracted the action potentials of single-unit activity from raw MER recording and remove the mean to get rid of the power at  $f=0$  Hz. In the following step, we removed the power line noise (50 Hz) and its first five harmonics based on spectrum interpolation technique to prevent introducing bias to our power spectrum and neural oscillation estimation[42]. We adopted the Welch PSD estimation by selecting the Hanning window as our tapering function with a window size equal to the data length divided by 20. The overlapping of the windows is selected as 0.5. In the last step of power spectrum estimation, we normalize the

power spectrum by dividing the power values in each frequency by the total power of the SUA. This normalization procedure is essential for acquiring a comparable power spectrum across all isolated single units.

### 2.6.5 Information analysis

Mutual information is a metric that endeavours to quantify the non-linear relationship between two variables. Briefly, it measures the decrease in entropy of one random variable by observing another random variable[43]. In this module (Figure 2, top, right box), we want to measure the relationship between candidate neural biomarkers and their position along with the two aforementioned reference frames (Section 2.6.2. and Figure 3). This entails computing the mutual information  $I(B,S)$  between the value of the biomarker and the position of the recording (1).

$$I(B,S) = \sum_{s \in S, b \in B} P(s) \cdot P(b|s) \log_2 \frac{P(b|s)}{P(b)} \quad (1)$$

Where  $B$  is the set of values of the biomarker and  $S$  is the set of recording positions (in terms of relative depth level for the 1D reference frame and MNI coordinates in 3D MNI reference frame localization scenarios).  $P(b)$  is the probability over all recording sites of the biomarker having value  $b$ . The continuous values of the neural biomarker( $b$ ) were binned into four levels and represented with their corresponding bin values.  $P(s)$  is the probability of having a recording from position  $s$  over all biomarker values (which, by design, is the inverse of the number of recording positions).  $P(b|s)$  is the conditional probability of observing the biomarker value of a neuron at spatial component  $s$ .

Probability estimates during mutual information calculation can be impacted in the case of a limited amount of data. Since information is computed as the difference between two entropies, entropy bias is a significant issue here. The difference between the two entropy values is generally quite smaller in both entropies. Hence, even a tiny proportionate bias error in the entropy can have a significant impact on the information. As a result, a bias correction procedure should be applied to the calculation of mutual information. We chose the Panzeri–Treves[44] strategy out of all the bias correction approaches to tackle the sampling bias difficulties, and we computed the mutual information for each candidate neural biomarker. The significance of computed mutual information was assessed by generating a null distribution with the bootstrap test with 500 iterations. Based on the z-score of the computed mutual information, we defined the significance value ( $z \geq 2$  for  $p \leq 0.05$ ). We used the Information Breakdown Toolbox[45] in MATLAB to compute all the aforementioned information values.

For the 1D reference frame, mutual information is defined along electrode trajectories, and we only selected biomarkers

that exhibit significant mutual information along these trajectories. Initially, we considered the depth level of neurons as the spatial location and the corresponding value of candidate neural biomarkers as the response for information analysis. Then we sorted and discretized the values of candidate biomarkers based on four equally populated bins. We had fifteen depth levels for the 1D reference frame.

In the 3D MNI reference frame, we determined the neuron locations based on matching the MER recording sites with their corresponding MNI coordinates (see Section 2.2.).

For utilizing the mutual information in 3D MNI space as a way of a feature reduction for our neural decoders, similar to the 1D reference frame, we adopted the approach of Mosher *et al*[19]. As opposed to their work, we did not employ principal component analysis (PCA) for feature reduction. We translated their optimal axis description to our preferential directions of candidate neural biomarkers definition. Instead of correlating the values of candidate biomarkers on rotated axes with anatomical positions of neurons, we used the mutual information to also grasp more complex non-linear relationships between the value distributions of candidate biomarkers and the MNI coordinates of neurons.

To define the preferential direction, i.e., the direction with the highest significant mutual information, of candidate neural biomarkers, we scanned the whole GPi. We defined two initial reference lines for the GPi scanning process. For the right hemisphere, we determined a line passing through the average target point of GPi DBS for the right hemisphere and the most lateral & anterior tip of GPi. Along this line, we define two tip points  $P_A$  and  $P_B$  that has a length three times the distance between  $GPi_{tip}$  and  $GPi_{target}$  points (Figure 3A). Similarly, for the left hemisphere, we defined another line passing through the target point of GPi DBS and the most medial & posterior tip of the left hemisphere GPi. Again, we defined two tip points  $P_A$  and  $P_B$  (Figure 3B). We rotated these initial lines with Euler angles from 0 to 180 degrees by 30-degree steps in the x, y, and z axes. Hence, this scanning procedure resulted in 343 rotated directions inside GPi (Figure 3C). On each of these 343 lines, we projected the neurons according to their MNI locations. Following the projection procedure, we split each line into 15 equally sized sections due to having a similar number of divisions with depth levels of electrode trajectories in the 1D reference frame. Therefore, the section on the selected direction that coincides with the projection of neuron location is used for mutual information calculation with the biomarker values of neurons in the 3D MNI reference frame.

To be more rigorous for the localization tasks that occur in a specific anatomical axis of GPi (i.e., dorsoventral localization of neurons within GPi), we did not consider mutual information results in all defined directions. Instead,

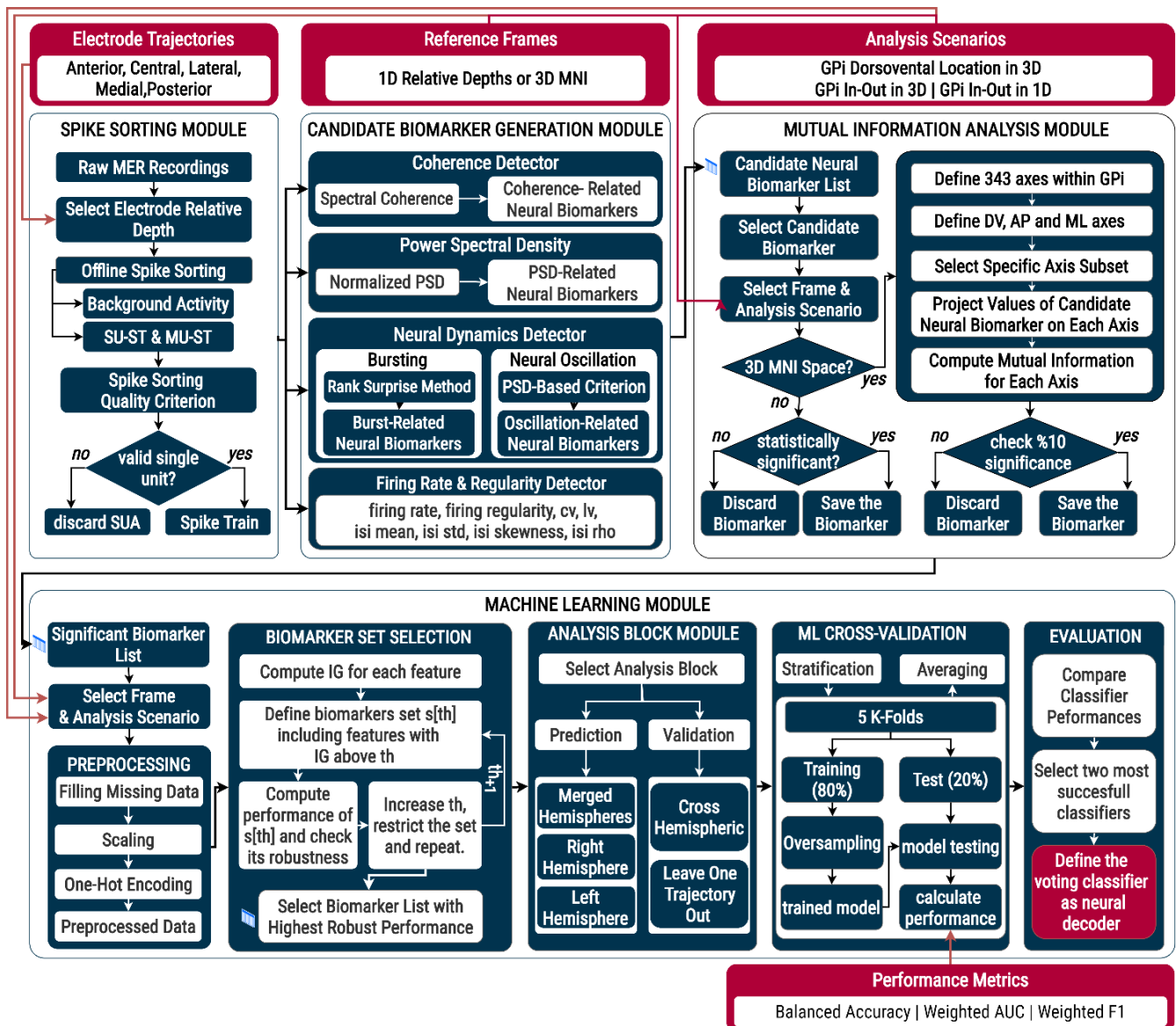
we grouped these predefined 343 directions into three subsets. To achieve this division, we calculated the angle between each direction and all three anatomical axes that have the origin at the  $GPI_{target}$  point. Then we assigned each direction to the anatomical axis that has the smallest angle with it. With this approach, we defined subsets of directions that are aligned mostly with a specific anatomical axis of GPI. The red(N=112), green(N=111), and blue(N=120) coloured directions in Figure 3C indicate the subsets of direction that belong to the anteroposterior, dorsoventral, and mediolateral axes of GPI respectively. We utilized these three groups as follows: for the dorsoventral localization task, we only considered candidate neural biomarkers that

show significant mutual information in at least 10% of dorsoventral directions. Thus, we evaluated the mutual information only in the relevant anatomical axis for each localization task and ignored the remaining directions.

We exploited the results of mutual information analysis as the first step of the feature selection mechanism for the machine learning module. The detailed information regarding the mutual information computation for different tasks can be found in the supplementary material.

### 2.6.6 Machine learning decoding algorithm

The last module of the computational pipeline (Figure 2,



**Figure 2. Modular structure of neural decoder pipeline for neuron localization.** The computational pipeline of the single-unit activity localization is based on the use of candidate neural biomarker definitions in machine learning algorithms. The pipeline is made up of four different modules: Spike Sorting Module, Candidate Neural Biomarker Generation Module, Mutual Information Analysis Module, and Machine Learning Module. Abbreviations: AP, anteroposterior; DV, dorsoventral; IG, information gain; ML, mediolateral.

bottom) is the machine learning module that contains the implementation of multiple neural decoders and cross-validation approaches to localize neurons in two different reference frames (1D and 3D). Apart from different reference frame definitions, we specified two analysis blocks: prediction and validation blocks. In terms of the prediction block, we conducted merged-hemispheres, right-hemisphere-only, and left-hemisphere-only classification tasks. To cross-validate the performances of trained models in the prediction block, we developed seven cross-validation classification tasks in the validation block. Firstly, we implemented five

leave-one-trajectory-out validation tasks to check the performance of the merged-hemispheres model for each trajectory separately. We trained the merged-hemisphere model with neurons of four trajectories and completed the testing phase with the neurons of the remaining trajectory. For cross-hemispheric validation tasks, we tested the performances of trained right-hemisphere and left-hemisphere models using neurons of the opposite hemisphere.

We have started our machine learning procedure with the classical preprocessing steps: filling the missing values, feature scaling, and standardization[46]. Following these preliminary steps, we defined additional feature selection criteria to our initial mutual information-based approach. As mentioned previously, the work of Mosher et al. exploited the principal component definitions for feature reduction. Following the mutual information-based first feature selection, we needed to conduct an additional step in some cases to further remove candidate neural biomarkers. For example, some candidate neural biomarkers showed isotropic behaviour in terms of mutual information in 3D MNI space (see Supplementary Material Figure 4). Isotropic candidate biomarkers are not particularly suitable to investigate the electrophysiological structure of dystonic GPi. Thus, some of the candidate biomarkers passed the mutual information-based feature reduction but decreased the overall performance of 3D localization tasks. To pursue the use of information theory as our main feature selection mechanism, we adopted an entropy-based approach. We constructed a Random Forest[47] model with the information gain[48] metric which is an index that measures the entropy of the candidate neural biomarker for the classification task to define its importance. Using the same model, we calculated the capability of each candidate neural biomarker concerning the amount of entropy it can describe. We defined then a set of subsets of candidate biomarkers including only those with information gain above iteratively increasing threshold values, and for each of these subsets, we computed the localization performance and its robustness (see Supplementary Materials 6.1. for details). We finally selected the subset with the optimal and robust performance. This

final step is hence based on the biomarker set performance, rather than on the individual performance of each biomarker.

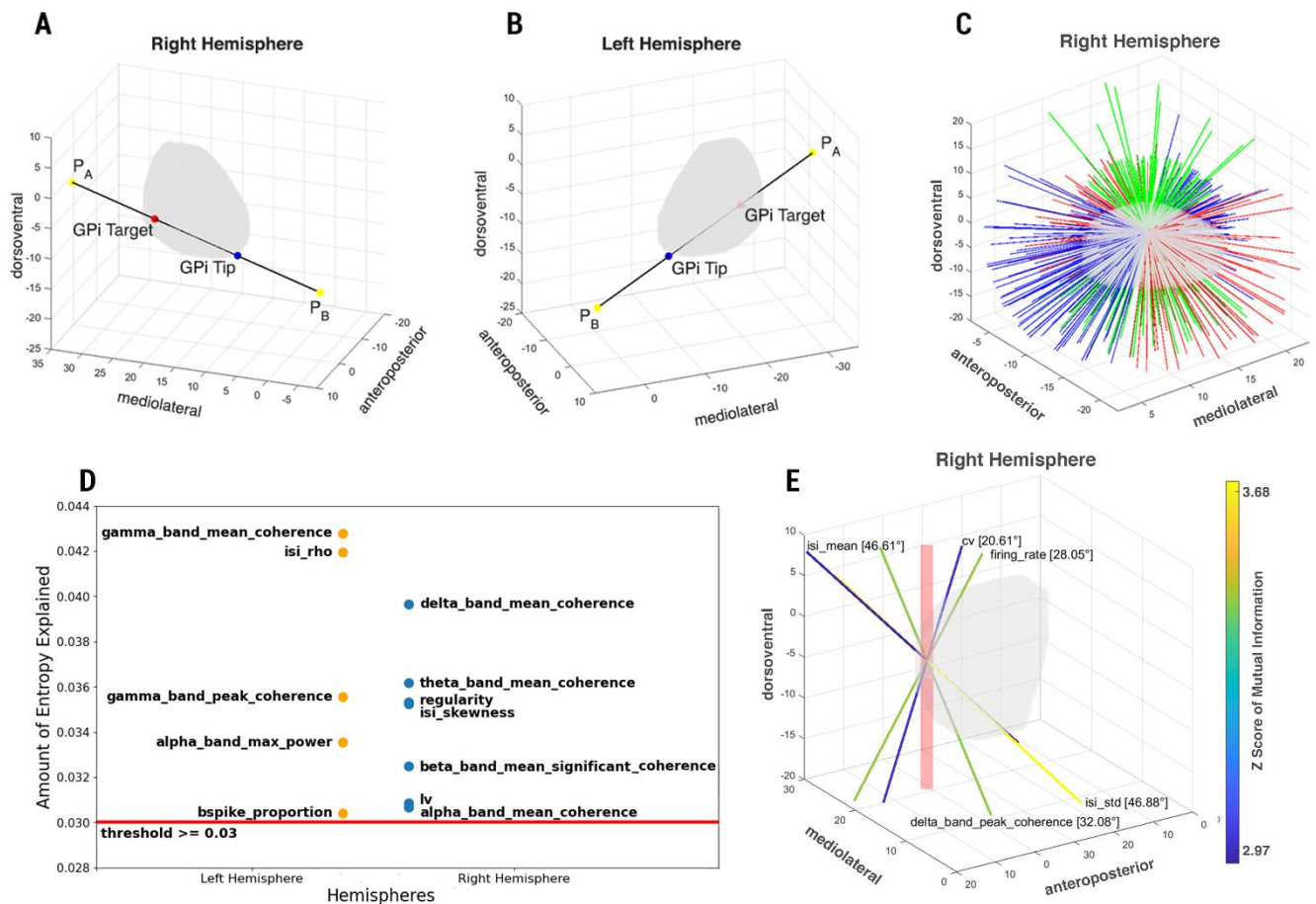
We repeated this two-level feature selection approach for all tasks in the prediction block to select for each task the optimal set of biomarkers. For the classification process, we have selected Decision Tree[49], Random Forest[47,49], K-Nearest Neighbors[49], Gaussian Process Classifier[49], and Support Vector Machine[49] as our main classifiers. Because of the class imbalance, we assessed classification performance by computing balanced accuracy, weighted AUC (area under the curve) score, and weighted F1 score. In the training and testing phases of our machine learning models, we adopted the Stratified K-Fold Cross Validation[49] procedure to prevent possible biases of classical training and test split approach. In each split of five-folds[50], we applied SMOTE[49] oversampling with additive white gaussian noise (AWGN) (standard deviation = 0.005) for overcoming imbalanced data problems in the training set. In each fold, we calculated our performance metrics with the test set and averaged scores across all five folds to acquire the mean performance. As a final step in this module, we compared the performances of all five neural decoders across all tasks in the prediction block and selected the best two classifiers for the Voting classifier with a soft voting approach. With the Voting classifier, we tried to further improve the localization performances by generating our unique ensemble model. Instead of using all five trained models, we only selected the best performing two models because, for some classification tasks, we experienced degraded performances in some of the trained classifiers. Hence, including these classifiers in Voting Classifiers reduced the overall performance.

## Results

### 3.1 Feature Selection

Before presenting the results of our two-step feature selection procedure, we wanted to emphasize an important fact. The fraction of neurons displaying oscillatory behaviour is quite low (see Table 1). Since neural oscillations are not significantly evident for dystonic GPi neurons in MERs at the single unit activity level, we did not come across any oscillation-related candidate neural biomarkers to pass the feature selection steps and to be included in the neural decoder in any type of localization.

We employed the information theory paradigm to generate significant neural biomarker sets from all candidate neural biomarkers, as indicated in the Methods section. Hence, we were able to achieve a cohesive two feature selection steps with exploding two different metrics: mutual information and information gain in the first and second steps respectively to boost neural decoder performances.



**Figure 3. The initial direction definitions in both hemispheres and the lines scanning right hemisphere GPI for mutual information calculation in MNI space.** (A) The initial line in the right hemisphere GPI that pass through  $GPI_{tip}$  and  $GPI_{target}$  and lies between  $P_A$  and  $P_B$  tips. (B) The initial line in the left hemisphere GPI that pass through  $GPI_{tip}$  and  $GPI_{target}$  and lies between  $P_A$  and  $P_B$  tips. (C) Directions that are generated based on the rotation of the initial line by Euler angles with 30-degrees steps in x, y and z axes to scan the whole GPI. The colour of directions represents the directions that belong to a specific anatomical axis. (D) The significant neural biomarkers for GPI in-out classification in 3D MNI reference frame for right and left hemisphere localization tasks in prediction block (mean explained entropy threshold  $\geq 0.03$ ). (E) The preferential directions of significant five neural biomarkers for dorsoventral localization in 3D MNI reference frame for merged hemispheres task of prediction block. The numeric values denote the angle between the preferential direction of the neural biomarkers and the dorsoventral anatomical axis of GPI. The red column represents the main direction of electrode trajectories that was used during GPI DBS surgeries.

Following our feature selection process in the GPI in-out localization task in the 1D reference frame, we ended up with seven significant neural biomarkers: two firing regularity-related and five power spectral density-related neural biomarkers (green rows in Table 2).

Contrary to other localization tasks, the mutual information-based feature selection step did not eliminate any candidate neural biomarkers for GPI in-out localization in the 3D MNI reference frame. The other analysis scenarios occurred mainly in one direction independently from the

Neural Dynamic	Total Frequency	Frequency in Left Hemisphere	Frequency in Right hemisphere
delta band oscillation	0.453%	0.571%	0.321%
theta band oscillation	1.057%	0.857%	1.282%
alpha band oscillation	0.453%	0.571%	0.321%
beta band oscillation	3.172%	2%	4.487%
gamma band oscillation	16.616%	12.857%	20.833%

**Table 1. The oscillatory behaviour of single unit activities.** The existence of significant neural oscillations in dystonic GPI neurons are quite low. Only the gamma band oscillations are comparably more common with slightly higher than 15%.



Feature	Feature Selection 1st Step				Feature Selection 2nd Step
	Mutual Information	Z-Score	Bootstrap Mean	Bootstrap Standard Deviation	Entropy Described by Neural Biomarker
firing_rate	0.03924	3.53923	0.00311	0.01021	0.042093
cv	0.03383	2.92200	0.00362	0.01034	0.062429
isi_mean	0.03967	3.49479	0.00345	0.01036	0.042329
isi_std	0.04272	3.87863	0.00393	0.01000	0.057545
intrabf	0.03354	2.61559	0.00682	0.01022	0.017214
interbi	0.04765	4.00772	0.00675	0.01021	0.015370
burst_avg_spikes	0.03178	2.22446	0.00835	0.01053	0.016160
delta_band_mean_power	0.05476	5.03064	0.00306	0.01028	0.041697
delta_band_max_power	0.04259	3.58223	0.00359	0.01089	0.050061
delta_band_min_power	0.04786	4.26317	0.00363	0.01038	0.050359
theta_band_mean_power	0.04478	3.90063	0.00414	0.01042	0.041190
theta_band_max_power	0.05748	5.22345	0.00380	0.01028	0.043302
theta_band_min_power	0.05551	5.01112	0.00377	0.01032	0.051437
alpha_band_mean_power	0.05406	4.93120	0.00376	0.01020	0.037910
alpha_band_max_power	0.04472	3.90169	0.00358	0.01055	0.038359
alpha_band_min_power	0.07433	6.73970	0.00365	0.01049	0.050350
beta_band_mean_power	0.06395	5.95098	0.00380	0.01011	0.042108
beta_band_max_power	0.04703	4.27481	0.00316	0.01026	0.042381
beta_band_min_power	0.05765	5.17704	0.00332	0.01049	0.042063
gamma_band_mean_power	0.04945	4.42030	0.00373	0.01034	0.043737
gamma_band_max_power	0.03778	3.25407	0.00370	0.01047	0.043269
gamma_band_min_power	0.05880	5.03240	0.00354	0.01098	0.051960
delta_band_peak_coherence	0.02870	2.34981	0.00439	0.01035	0.022294
delta_band_peak_coherence_frequency	0.02362	1.78176	0.00523	0.01032	0.027913
theta_band_mean_significant_coherence	0.02613	1.94223	0.00608	0.01032	0.026470

**Table 2. Two-step neural biomarker selection results of GPI in-out neuron localization for merged hemispheres task in the 1D relative depth reference frame.** The table presents the results of the feature selection process of GPI in-out localization in 1D reference frame. Only features that show significant mutual information in the first step of feature selection are presented in this table. Green highlighted rows indicate biomarkers selected by the process for this task.

reference frames. But for the GPI in-out task in the MNI reference frame, we were considering the 3D structure of the GPI in all directions. Since we were only interested in whether the neuron is located inside GPI, determining mutual information in a specific axis loses relevancy. The other main difference is that we decreased our mean entropy threshold from 0.05 to 0.03 (again a heuristic threshold) (see Supplementary Material Section 6.1.) considering the higher number (no elimination of features previously) of markers evaluated for mean explained entropy. Hence, it got harder for a biomarker to explain the fraction of total entropy by 0.05 individually. We defined two final sets of significant neural biomarkers. The correlation coefficient of ISI distribution, the proportion of spike lies in bursting intervals, the maximum power in the alpha band, the mean and peak coherence in the gamma band was selected for the GPI in-out localization in the left hemisphere. For the right hemisphere, firing regularity, local variation, ISI distribution skewness, the significant portion of mean coherence in the delta band, and mean coherences delta, theta, and alpha frequency bands

provided entropy decrease that surpasses the predefined threshold value (see Figure 3D, Supplementary Material Table 8).

In terms of dorsoventral localization in the 3D MNI reference frame, we identified significant neural markers carrying information about the spatial structure of dystonic GPI (see Method, Figure 3E and Figure 4A) for all three tasks in the prediction block. Oscillation-related neural biomarkers were not informative at all in the dorsoventral direction with no biomarker passing the 10% threshold in the first step of our feature selection. In terms of bursting-related markers, 4 out of 6 of them carried some level of information in the dorsoventral axis, but they failed to decrease the entropy of the dorsoventral localization problem adequately (see Supplementary Material Table 9). Some firing rate and regularity-related biomarkers (mean firing rate, the standard deviation of ISI, and coefficient of variation) were generally essential for dorsoventral localization tasks in the prediction block (see Methods and Figure 7B). ISI mean was also labelled as a descriptive marker for the dorsoventral

localization tasks in merged and right hemispheres cases. Moreover, we found firing rate, ISI mean and ISI standard deviation for merged hemispheres, cv and ISI standard deviation for left hemisphere GPi were significantly different between dorsal and ventral neurons (the median split: -3.880 on the Z-axis of GPi definition in DISTAL Atlas,  $p < 0.05$  Holm-Bonferroni correction applied for Mann-Whitney U test after Cramer-Von Mises test, see Supplementary Material Table 18 and Figure 7A).

### 3.2 GPi in-out Localization in 1D Reference Frame

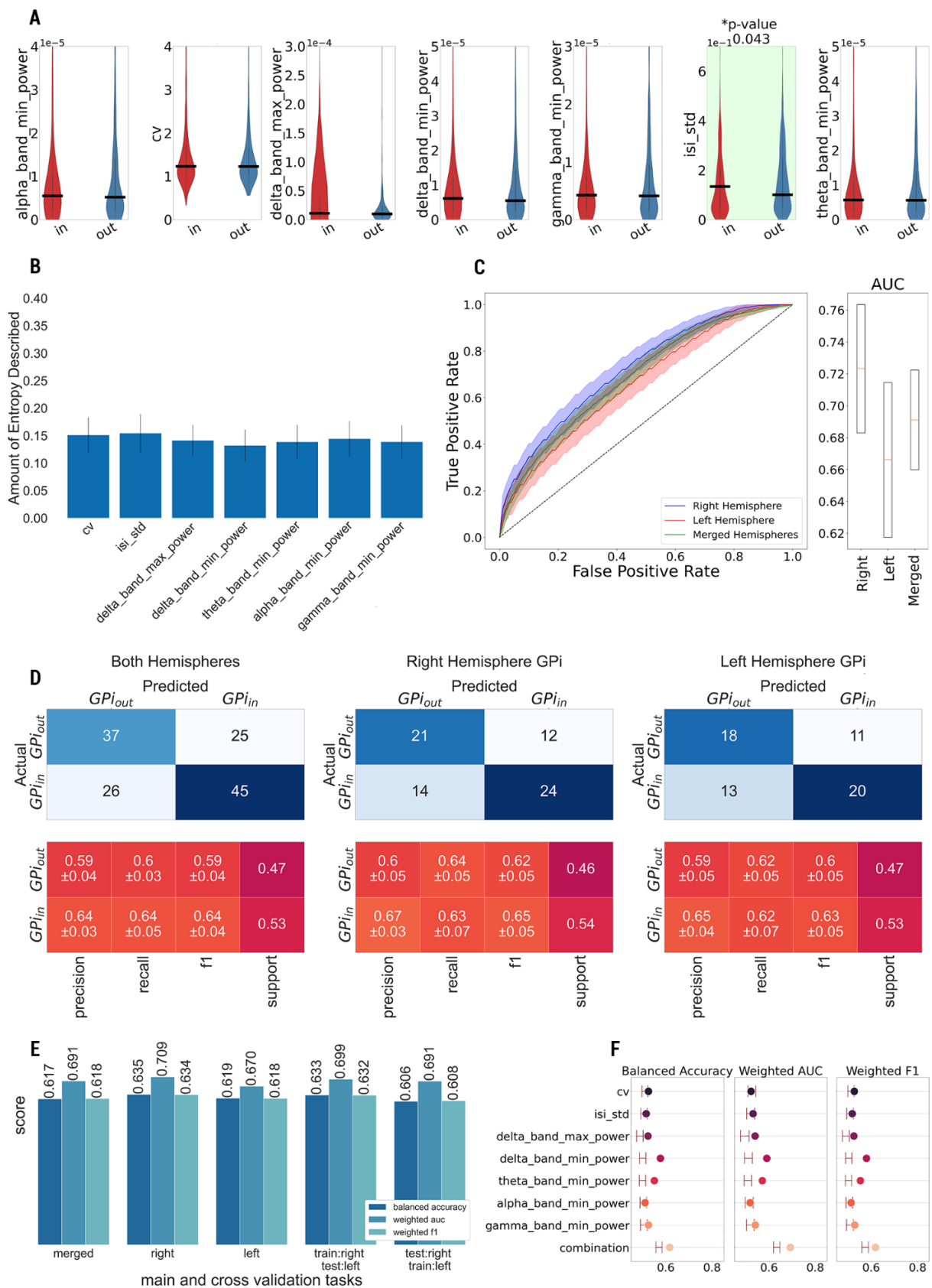
In the first classification task, we aimed to localize neurons in terms of their positioning whether inside or outside of GPi in the 1D reference frame across trajectories (see Methods). Based on the intraoperative notes of a neurophysiological expert, we labelled the neurons on trajectories in terms of their location. GPi-in neurons are defined as neurons that are positioned between -4mm and -10mm from the target location. The remaining neurons were considered to be outside of the GPi.

MERs are performed during DBS implant to support the localization of the target area after the initial guess on its position is performed with pre-surgical imaging. Hence, before moving to use MER data to analyse the difference in the activity over the volume of the GPi, we assessed which of the neural markers (see Methods) could carry more information about the position of a given recording site within or outside the GPi.

Following our feature selection process, we ended up with seven significant neural biomarkers: two firing regularity and five power spectral density-related neural biomarkers (green rows in Table 2). The standard deviation of the ISI distribution was the only neural biomarker significantly differentiated inside and outside the GPi ( $p < 0.05$ , Mann-Whitney U test, with Holm-Bonferroni correction) (see Figure 4A, Supplementary Material Table 10). However, in terms of the amount of entropy they can explain, neural biomarkers are relatively similar (see Figure 4B). Individual performances of trained machine learning models were evaluated with the Stratified 5-Fold cross-validation with training set upsampling (see Methods). The best performance was achieved with the KNN classifier with 0.623 mean balanced accuracy, 0.692 mean weighted AUC score, and 0.624 mean weighted F1 scores across all five folds. Random Forest Tree algorithm, the second-best performing model, provided a slightly worse performance with a 0.619 mean weighted AUC score. We used both KNN and Random Forest models in an ensemble model to generate a Voting Classifier with the soft voting mechanism (Supplementary Material Table 11). To measure the robustness of the Voting Classifier, we applied a bootstrap test with 200 iterations (each iteration has also Stratified 5-Fold cross-validation). We acquired the ROC curves across all iterations and

generated the mean ROC curve with the 95% percentile confidence interval around it which is given in Figure 4C. This classification was performed for each hemisphere separately and then for the merged hemispheres. Based on the mean ROC curve, the Voting classifier trained for left hemisphere GPi has the best performance followed by merged hemisphere and right-hemisphere classifiers. Apart from the ROC curves, we also illustrated the mean AUC scores and their corresponding confidence interval with 95% percentiles in the right panel of Figure 4C, and the performance differences are also evident in terms of AUC scores, although with similar ranges for confidence intervals of mean AUC score. Note that the class imbalance is not a big problem for GPi in-out classification in the 1D reference frame since we have almost 47% of neurons situated inside of the GPi and 53% remaining located outside of the GPi (see Figure 4D). The performance similarities across all tasks in prediction and validation blocks are also notable in this type of localization. The mean precision, recall, and accuracy scores range in a limited region from 0.59 to 0.65 across all prediction block tasks. Mean balanced accuracy, mean weighted AUC score, and mean weighted F1 scores of all tasks in prediction and two cross-hemispheric validation tasks are given in Figure 4E. In all tasks in the prediction block, we observed approximately 0.7 mean weighted AUC scores. The remaining two performance metrics were also quite similar. It is notable that, even though we trained Voting classifiers with the neurons of one hemisphere and test them with the neurons of the opposite hemisphere, the performance is again similar. Hence, the hemispheric effect for dystonic GPi to decide whether the neuron is inside or outside of the GPi is not crucial for the 1D reference frame.

Further, we assessed the localization capabilities of selected neural biomarkers individually by comparing them to the combined version, the ideal neural decoder. To achieve this final analysis, we trained the previously generated Voting Classifiers with only a single neural biomarker and computed the mean of performance metrics across five stratified folds, with the same bootstrap ( $N=200$ ) estimation of the performance described above. Individual performances of the coefficient of variation, the standard deviation of ISI distribution, and the minimum power in alpha and beta bands occurred within the random chance intervals (Figure 4F). The minimum power in the delta band showed the best individual performances across all seven selected neural biomarkers. No neural biomarkers delivered a higher or comparable result similar to the combined version of the biomarkers. Overall, the results were robust and significant, but the performance was far from satisfying and required the combination of multiple markers. We concluded that the ability of MER in identifying GPi boundaries by performing the analysis in the 1D space given by penetration tracks is limited in our dataset.



**Figure 4. The results of GPI in-out neuron localization in the 1D relative depth reference frame.** (A) The value distribution of significant neural biomarkers for GPI-in and GPI-out neural populations and their corresponding probability values (corrected with Holm-Bonferroni Method) of Mann-Whitney U tests. (B) Significance of neural biomarkers in terms of their abilities to explain mean entropy during the generation of Random Forest Trees classifier (The values are computed again after the completion of the feature selection process with only significant neural biomarkers). (C) The ROC curves of the Voting Classifiers (constructed with trained Random Forest Trees and KNN classifiers) for different tasks of prediction block with their corresponding percentile-based confidence intervals (95%) of AUC scores. (D) The confusion matrices of GPI in-out classification tasks in right-hemisphere-only, left-hemisphere-only, and merged-hemispheres tasks and their classification performances (precision, recall, and F1 scores). (E) The performances of the Voting Classifier models in respect of balanced accuracy, weighted AUC, and weighted F1 scores in tasks of prediction and validation blocks. (F) The performance of the Voting Classifier model for single and combined versions of the significant neural biomarkers in merged hemisphere tasks. The percentile-based confidence interval of the random chance scenario for each performance metric is indicated with two horizontal lines (left whisker: lower bound, right whisker: upper bound of confidence interval).

### 3.3 GPI in-out Localization in 3D MNI Reference Frame

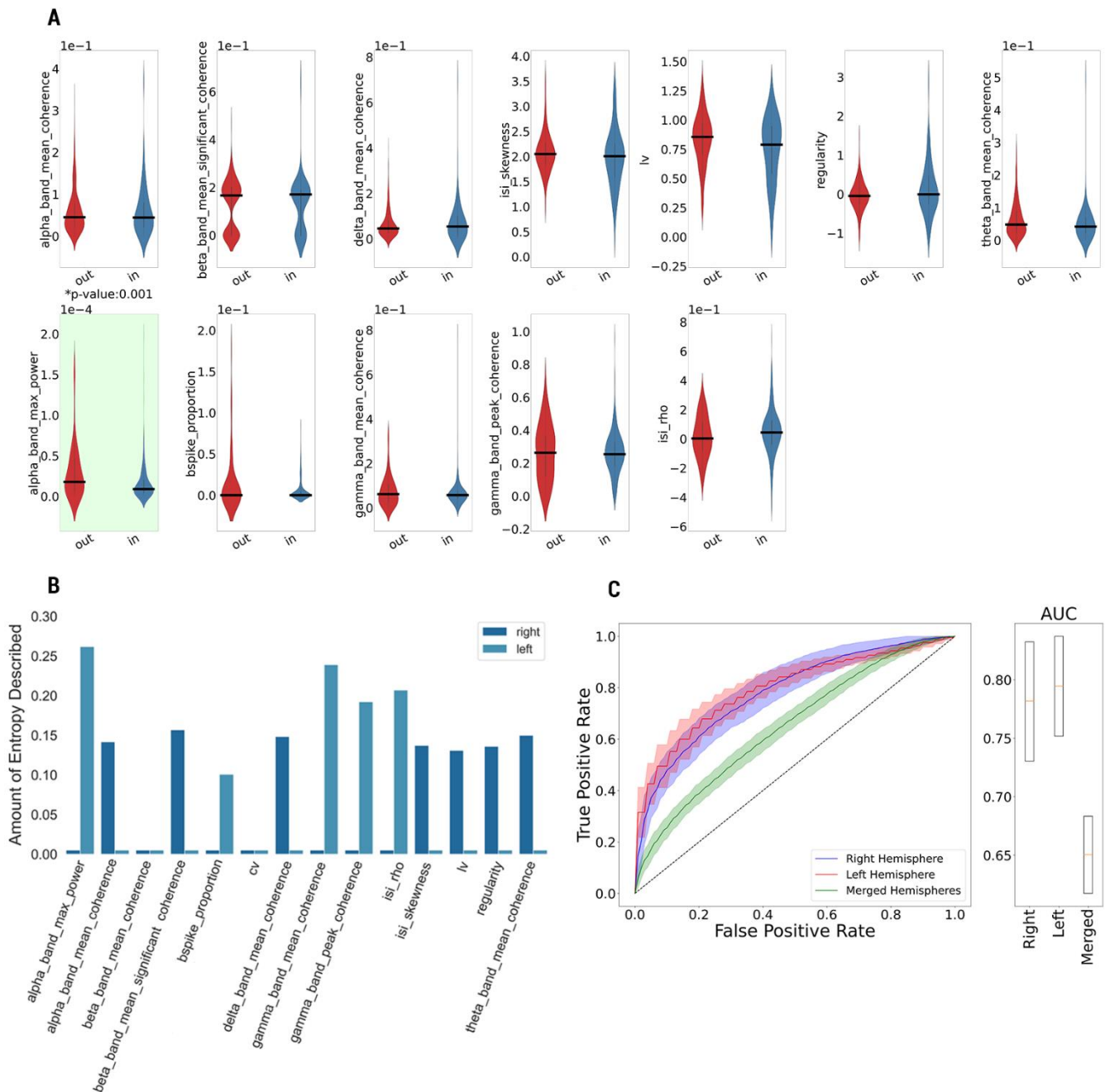
To localize the neurons based on the 3D boundaries of GPI, we resorted to the 3D distribution of firing properties in MNI space (see Methods). Briefly, we designed a similar neuron localization analysis in MNI space considering the anatomical structure of GPI using the DISTAL Atlas. To establish the borders of GPI in the MNI reference frame, we used a 0.5 probabilistic value as our heuristic threshold for the DISTAL atlas. This section will provide the complete results of the individual hemispheric tasks in the prediction block of this localization task. Indeed, the performance of GPI in-out localization is significantly decreased by merging the hemispheres (see Supplementary Material Table 15).

Initially, we labelled all neurons as GPI-in or GPI-out in both hemispheres. As previously indicated, we leverage the same computational pipeline for all localization tasks to have comparable results. We initiated our second localization task with feature selection again.

Only, the maximum power in the alpha band significantly differentiated in left hemisphere GPI (p-value = 0.0011 after the Mann-Whitney U test with Holm-Bonferroni correction) (see Supplementary Material Table 14, Figure 5A, the top panel). Contrary to the previous localization task, we did not have a uniform entropy behaviour across all markers. The entropy behaviour of seven markers in the right hemisphere was considerably uniform with a value of 0.15. Contradictorily, we had a more uneven explained entropy distribution for the left hemisphere (see Figure 5B). As mentioned above, we had huge performance differences between merged and split hemisphere tasks in the prediction block. The mean weighted AUC scores reached 0.78 for both the right and left hemisphere GPI, meanwhile, it stayed at 0.631 in the merged hemisphere case. The performance of the trained neural decoder further decreased in the cross-hemispheric cross-validation tasks reaching below 0.62 mean weighted AUC score (see Figure 6B). Similar to selected performance metrics, the behaviour of acquired ROC curves in the bootstrap test with 200 iterations shows parallelism. We acquired mean AUC scores of 0.773 and 0.792 for the right and left hemispheres respectively (see Figure 5C, the right panel). Even the upper limit of the 95% confidence

interval of the merged hemisphere neural decoder was not able to surpass 0.70 AUC scores. This situation can be explained for multiple reasons. Initially, when we tried to measure population-wise median differences, since none of the neural biomarkers follows the normal distribution, no marker showed a statistically significant difference between the two neural populations in the merged-hemisphere task. We further investigated to reveal the root cause of performance degradation as follows: we compared biomarkers for GPI-in and GPI-out populations of opposite hemispheres to see if we had any significant differences. Surprisingly, power spectral density-related fifteen biomarkers differed significantly between GPI-in and GPI-out populations of both hemispheres ( $\alpha=0.05$ ). Additionally, minimum power in the gamma band, peak coherence in the alpha band, and peak coherence in the gamma band were the three markers that showed significant population median differences for both GPI-in and GPI-out populations between the two hemispheres (see Supplementary Material). The third possible reason could be related to the drastically different class distribution in both hemispheres: 70(20%) GPI-in, 280(80%) GPI-out in the right hemisphere, 254(81%) GPI-in, and 54(19%) GPI-out neurons. Notably, the class distribution was exactly the opposite in both hemispheres. Even though the merger of hemispheres for complete analysis had an almost balanced class distribution (49% GPI-in, 51% GPI-out), it did not provide better localization performance. The performance superiority of split hemisphere scenarios also was supported by the capabilities of classifiers to label majority class neurons correctly with moderate success in the minority neuron classes (see Figure 6A).

In terms of individual capabilities of biomarkers in both hemispheres, they were not able to provide a better mean balanced accuracy performance than the 95% confidence interval of random chance (see Figure 6C). The high F1 scores of individual markers and confidence interval of random chances that were constructed with F1 are mostly the side effect of keeping the class distribution during the random chance bootstrap test. In the worst-case scenario, 60% of majority class neurons will be labelled the same since the minority neural population is only made up of 20% of all

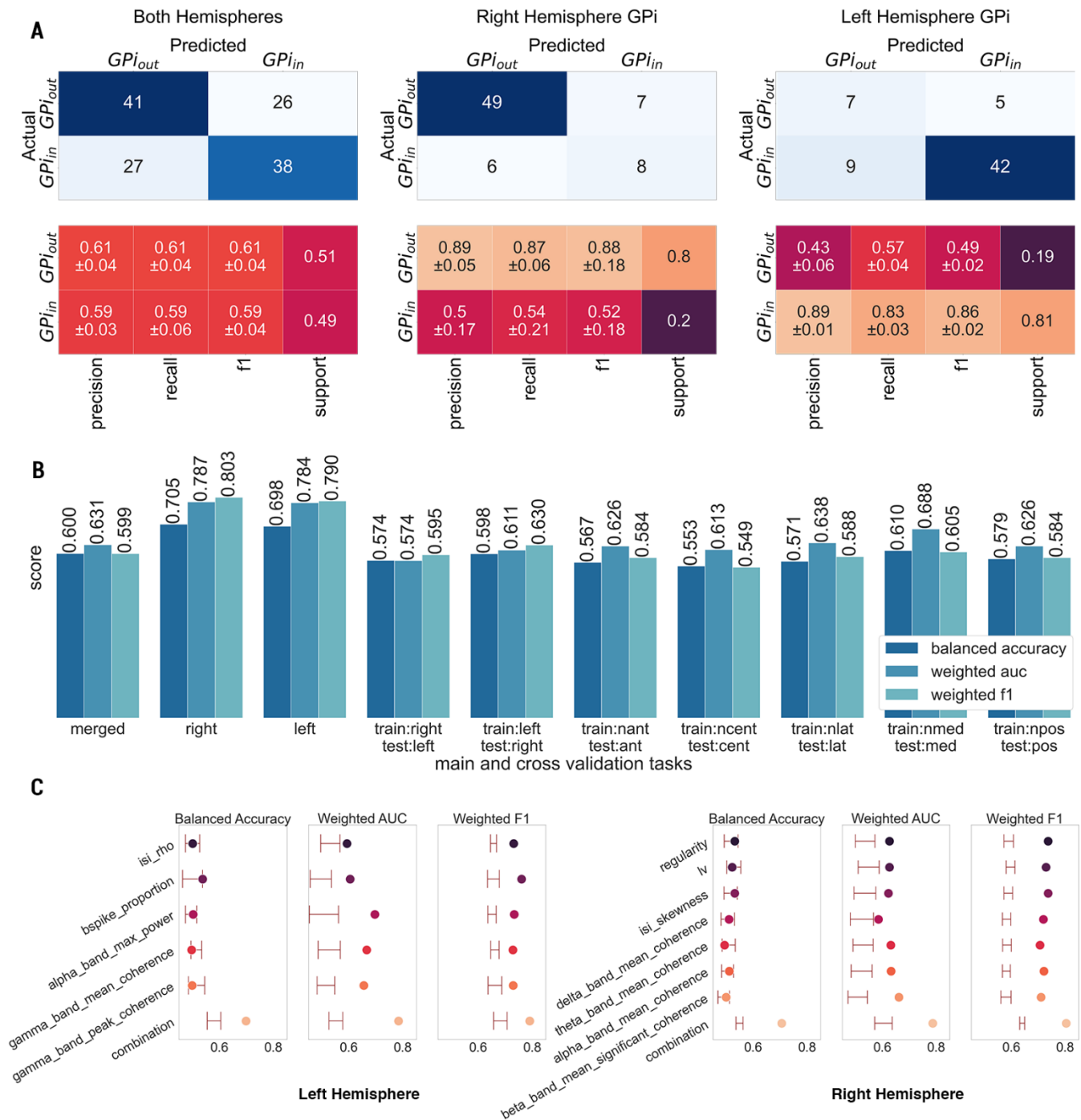


**Figure 5- The results of GPI in-out neuron localization in the 3D MNI reference frame. (A)** The value distribution of significant neural biomarkers for GPI-in and GPI-out neural populations and the corresponding probability values (corrected with Holm-Bonferroni Method) of Mann-Whitney U tests (top: right hemisphere neurons, bottom: left hemisphere neurons). **(B)** Significance of neural biomarkers in terms of their abilities to explain mean entropy during the generation of Random Forest Trees classifier for right hemisphere and left hemisphere neurons individually. **(C)** The ROC curves of the Voting Classifiers (constructed with trained Random Forest Trees and Gaussian Process classifiers) for different analysis scenarios with their corresponding percentile-based confidence intervals of AUC scores.

neurons in both hemispheres. The power of the combined version of neural biomarkers for both hemispheres was especially more evident in this neural localization task.

### 3.4 Dorsoventral Localization of Spike Dynamics in GPI

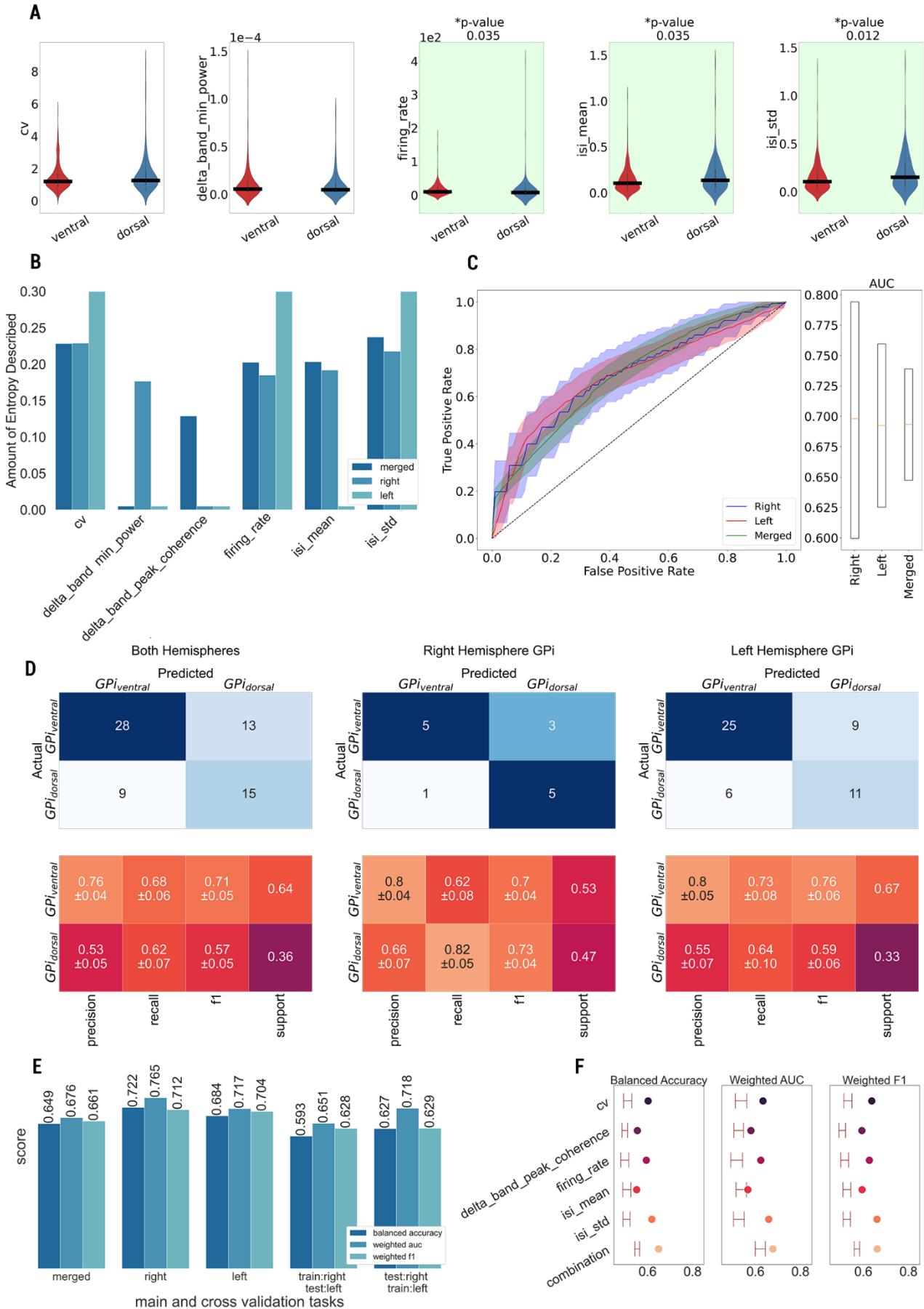
We next investigated whether the neural dynamics identified so far differ across the spatial structure of the dystonic GPI (in particular, along the vertical dorsoventral axis). Akin to GPI in-out localization in the 1D reference frame, we will present the complete analysis results of the merged hemisphere task of the prediction block.



**Figure 6- The results of GPI in-out neuron localization in the 3D MNI space domain. (A)** The confusion matrices of GPI in-out classification tasks in right-hemisphere-only, left-hemisphere-only, and merged-hemispheres analysis modes and their classification performances (precision, recall, and F1 scores) in MNI space. **(B)** Performance of the Voting Classifier models using the balanced accuracy, weighted AUC, and weighted F1 scores in tasks of prediction and validation blocks. (trajectories: and(anterior), cent(central), lat(lateral), med(medial), pos(posterior), nant(non-anterior), ncent(non-central), nlat(non-lateral), nmed(non-medial), npos(non-posterior)). **(C)** The performance of the Voting Classifier model for single and combined versions of the significant neural biomarkers in merged hemisphere analysis scenario. The percentile-based confidence interval of the random chance scenario for each performance metric is indicated with two horizontal lines (left whisker: lower bound, right whisker: upper bound of confidence interval).

The discrimination capabilities of some of the individual markers (the standard deviation of ISI distribution and coefficient of variation) were close to the combined version of biomarkers and way higher than the chance within the

95% confidence range (see Figure 7F). However, the combined version of neural biomarkers for Voting Classifier (KNN with 0.628 mean balanced accuracy, 0.690 mean weighted AUC score, 0.638 mean weighted F1 score, and RF



**Figure 7- The results of dorsoventral neuron localization within GPI in the 3D MNI space domain.** (A) The value distribution of significant neural biomarkers for dorsally and ventrally located neural populations and the corresponding probability values (corrected with Holm-Bonferroni Method) of Mann-Whitney U tests in merged hemispheres analysis task. (B) Significance of neural biomarkers in terms of their abilities to explain mean entropy during the generation of Random Forest Trees classifier for dorsoventral localization task in all tasks in prediction block. (C) The ROC curves of the Voting Classifiers (constructed with trained Random Forest Trees and KNN classifiers) for different analysis scenarios with their corresponding percentile-based confidence intervals of AUC scores. (D) The confusion matrices of GPI dorsoventral classification tasks in prediction block tasks and their classification performances (precision, recall, and F1 scores) in MNI space. (E) The performances of the Voting Classifier models in respect of balanced accuracy, weighted AUC, and weighted F1 scores in tasks of prediction and validation blocks. (F) The performance of the Voting Classifier model for single and combined versions of the significant neural biomarkers in a merged hemisphere analysis scenario. The percentile-based confidence interval of the random chance scenario for each performance metric is indicated with two horizontal lines (left whisker: lower bound, right whisker: upper bound of confidence interval).

with 0.574 mean balanced accuracy, 0.602 mean weighted AUC score, 0.608 mean weighted F1 score) helped us to define the most robust neural decoder (see Supplementary Material Table 19). Our neural decoder has reached 0.649 mean balanced accuracy, 0.676 mean weighted AUC, and 0.661 mean weighted F1 scores and it performed significantly higher than 95% random chance confidence intervals (0.543, 0.565), (0.599, 0.644), (0.559, 0.583) respectively (see Figure 7F). The ROC curves of the main analysis tasks had a mean AUC score of 0.7. The only difference was the merged hemisphere task had the narrowest confidence interval with 95% confidence lying between 0.649 to 0.738 after 200 bootstrap iterations. (See Figure 7C, the right panel). Meanwhile, the right hemisphere decoder had a broader confidence interval since the right hemisphere drastically lowers the number of neurons located within GPI compared to the left hemispheres (70 and 254 respectively).

The class imbalance was more visible in the confusion matrices of prediction block tasks (see Figure 7D). We had class support for machine learning models 65% and 35% for ventrally and dorsally located neurons in merged hemispheres, and left hemisphere cases. Even though we had a more balanced neuron distribution in the right hemisphere, the low number of neurons created the main challenge in this task. The performance of neural decoders was relatively better when we conducted dorsoventral localization in right and left hemisphere GPI separately by a really small margin. The results of cross-hemispheric cross-validation tasks were comparatively lower than their prediction block counterparts. Especially performance decreased in the cross-validation task when the Voting Classifier is trained with 70 right hemisphere model and tested with 254 left hemisphere GPI neurons over stratified 5-fold cross-validation. The mean balanced accuracy dropped to 0.593 which is the lowest across all tasks in this localization procedure (see Figure 7E).

## Discussion

Our statistical analyses prove that some neural biomarkers related to specific spiking phenomena vary both in 1D relative depths and 3D MNI space. This is in striking contrast with the fact that GPI is considered a homogeneous subcortical structure. The results of this study imply that the

variety of neural dynamics inside dystonic GPI is high because we were able to define directions that showed statistically significant changes in terms of some neural biomarker values based on mutual information metrics. We used these results to locate neurons in three different localization tasks where we have comparable high classification performances for locating neurons based on different criteria with selected neural biomarkers. In our study, we try to collaborate on different neural biomarkers that are related to different spiking phenomena of dystonic GPI neurons. We can group these neural biomarkers into the following main categories: firing rate, firing irregularity, bursting, oscillation, spectral coherence, and power spectrum-related neural biomarkers. Even though we have found neural oscillations on MERs at the single neuron level, especially in the gamma band, frequency domain-associated neural biomarkers fail to describe the position of neurons in the dorsoventral localization task where we only consider neurons that are located inside the dystonic GPI. This result can be also interpreted from a neurological point of view. It may imply that the neural behaviour in terms of bursting, oscillation, and other spectral components of neural activity stay homogenous across the GPI at the single neuron level. Thus they are not reliable neural phenomena to locate the neurons inside the GPI. The other important fact, we have some significant frequency domain (oscillation, bursting, or coherence-based) neural biomarker that shows significance in terms of mutual information in the 1D relative depth frame. First and second-order statistics such as firing rate and firing regularity-related neural biomarkers vary significantly inside dystonic GPI.

It is important to note that our approach did not assume any a priori knowledge of the neurophysiological properties of the candidate neural biomarkers, as we only measured each biomarker's capacity to carry enough information to locate the DBS electrodes algorithmically within GPI. In future studies, we will address the underlying mechanism leading to spatial inhomogeneity of specific features of dystonic GPI activity.

### 4.1 Limitations



As cited earlier, we were not able to corroborate the anatomical position of the MERs electrodes with preoperative or postoperative medical images. Hence, the use of normalized probability-based anatomical atlas for the confirmation does not affirm the inter-subject anatomical variabilities of our subject group. The possibility of brain tissue displacements during GPi DBS surgery further added uncertainty to our findings. Furthermore, we believe that the GPi bounds were set conservatively by setting the heuristic threshold for the DISTAL atlas as 0.5. The number of neurons in the right GPi hemisphere is roughly 20%, and the majority of the electrode trajectories are outside of the right GPi. We know that before craniotomy, patients' heads are tilted at a precise angle. The tilt could potentially cause these types of misalignment between the real position of neurons and their DISTAL atlas-based location. Even though the use of a normalized atlas introduced some level of limitations, the DISTAL atlas is a widely adopted and validated atlas for DBS studies. In a recent study, MERs and DISTAL atlas were found to be consistent in determining the subthalamic nucleus (STN) entry for Parkinson's disease patients who underwent STN DBS surgery[18].

The expert definition was used for defining GPi borders in the 1D reference frame, as previously described. The concept introduces a level of uncertainty because we did not have relative locations of electrode trajectories to GPi.

We repeated the neuron localization procedure also relative to the mediolateral and anteroposterior axes. The main issue was we only had neurons located in the anterior and lateral sections of dystonic GPi in both hemispheres due to the location of trajectories of DBS electrodes. Hence, we were not able to discriminate neurons in all anatomical axes to get the complete functional structure of the dystonic GPi.

We are aware that working with SUAs that are isolated from the MER recordings poses methodological difficulties in terms of oscillatory behaviour detection, especially in single-trial experiments. Consequently, we do not assert that neural oscillations are not part of the pathophysiology of dystonia within the GPi. In our analysis, we only concluded that neuronal oscillations are not the optimal feature for lead localization based on MERs in GPi DBS surgeries for dystonia.

Finally, the acquired dataset was sufficient to lead to statistically robust results, mainly due to the high number of tracks per patient and the high quality of the recordings, but the set of patients was however limited to ten. Before translating to clinical application, larger datasets should be analysed to ensure that the proposed methods can overcome inter-subject differences.

#### 4.2 Algorithms performance

As we suggested in our introduction, the peculiar characteristics of the transition from the external segment of

the globus pallidus (GPe) to GPi, such as the sparseness of the neurons along the pallidum, the absence of a marked reduction in neuronal activity, and the high discharge rate variability, challenged the electrophysiological determination of GPi borders and led to AUC values of 0.67. Additionally, there was no well-established anatomical delineation between the dorsal and ventral regions of GPi for the dorsoventral location. As a result, we divided the GPi using the median split: -3.880 on the GPi Z-axis, which could also contribute to the performance of our pipeline. Hence, moderate values of AUC are less an indication of the efficacy of the method than a measure of the structural inhomogeneity of GPi. Apart from the neuroanatomical aspect of the GPi and the surrounding tissue, several factors (the issue regarding postoperative imaging, and conservatively defined GPi boundaries based on the Distal Atlas that we mentioned in our Limitation Section) introduced this uncertainty for the neuron location localization procedure, again leading to an AUC in the order of 0.67. However, cross-validation analyses and confidence interval estimation confirmed the robustness of the classification procedures.

We experimented with different classification approaches that exploit several mathematical concepts to distinguish particular neuron populations. Initially, we employed two tree-based classification algorithms: Decision Tree Classifier and Random Forest Classifiers. Both algorithms construct tree structures based on the capability of neural biomarkers to reduce impurity during the classification. The main advantage of the Random Forest Classifier is generating several random Decision Trees and conducting training instantaneously. The outcome of the classification process will be determined based on major votes between different trees. Being an ensemble approach, it generally provides robust results. For the K-Neighbour Classifier, the spatial structure of the dataset plays a primary role; definition of the distance metric, we used Euclidean distance, and the neighbour radius criteria can affect the classification results drastically. Considering that neural dynamics are entirely complex, the definition of apparent borders between neuron populations is not achievable for all localization procedures. Correspondingly, the spatial structure is vital for the classification of Support Vector Machines. Since SVM utilizes the definition of hyperplanes that separates the regions that belong to different classes, not all problems have a suitable structure for SVM. In the Gaussian Classifier case, the algorithm attempts to exploit the Gaussian probability distribution function similarly to Gaussian mixture problems. It tries to estimate the Gaussian distribution of each neural biomarker for each class in terms of mean and variance properties. Additional to these parameters, gaussian processes require specifying a kernel that manages how neurons relate to each other; specifically, it defines the covariance function of the data. The main problem might

occur when the distributions of two classes for a specific neural biomarker are noticeably similar. In this case, the algorithm cannot be able to discriminate neuron populations.

### 4.3 Lateralization

Our results posit that dystonic GPi may have a potential lateralization effect in terms of the in-out localization task conducted in the MNI reference frame. As stated in the result section, no neural biomarkers were selected for the neural decoders that were constructed for the right and left hemispheres individually. Furthermore, merging the neurons of two hemispheres onto the right GPi resulted in significant performance degradation.

The lateralization effect could be originated from two different sources. Firstly, the craniotomy for GPi DBS surgeries is initiated from the right side of the cranium of dystonic patients. Besides, when we examined the baseline motor indexes of the Burke-Fahn-Marsden Dystonia Rating Scale (BFMDRS) for the right arm, left arm, right leg, and left leg of patients, we discovered that the right sides of half of the patients are more affected by primary dystonia compared to their left sides (remaining 30% equally affected and 20% no baseline scores) (see Supplementary Material Table 22).

On the contrary, electrophysiological dynamics within GPi are consistent for two hemispheres, and the primary dystonia mostly modulates the mean firing rate and firing regularity of neurons.

### 4.4 Perspectives

We developed a method able to describe the functional structure of the GPi based on intraoperative multichannel high-density MERs. This approach can be fruitfully applied to find disease-specific and patient-specific partitioning on the GPi and hence support the localization of optimal DBS targets. Crucially, this method could also be applied to other neurological conditions to be treated by GPi implant, i.e., Tourette's syndrome and Parkinson's disease.

The proposed computational pipeline has the feasibility to operate in an online manner during DBS surgeries for lead localization. With a few straightforward changes to the pipeline, offline-to-online translation may be achieved. As an initial step, one of the recently proposed high-performance online spike sorting algorithms[51–53] can replace the offline counterpart that we utilized in our pipeline. Since each spike sorting algorithm adopts different computational and statistical approaches, the SUA generation time will differ. We approximated the computational running time of these online algorithms for 10-second recordings based on their reported benchmark tests. Mostly, they can extract single units in less than 1.5 seconds (starting from 0.64 to 1.49 seconds). Approximately  $1.25 \pm 0.24$  (follows the Gaussian distribution for 100 neurons) seconds are needed to

produce all the neural biomarkers we presented in the study for each neuron. Considering that the voting classifier was previously trained before the surgery, the localization of the neuron by the trained neural decoder takes around 0.98 seconds. With the aforementioned modifications, we thus think that an online implementation of our computational pipeline can predict the location of DBS leads in around 3.68 seconds (including the online spike sorting procedure) for each 10-second MER recording. In our dataset, we had an average of 10 depth levels per trajectory across 10 patients. Considering this fact, the total running time of the online version of the proposed pipeline takes around 37 seconds during the whole MER collection procedure for each trajectory. Additionally, the value distribution of significant neural biomarkers may also be illustrated, and the lead localization can be displayed on an anatomical atlas in real time. As a result, the pipeline we demonstrate here may serve as the basis for software that takes advantage of all the features and low computational time for providing feedback to neurosurgical teams in real-time.

### Ethical Approval

The study was approved by the Carlo Besta Institutional Ethics Committee (reference number 7/2014) and was carried out according to the Declaration of Helsinki. Informed written consent was obtained either from the patients or from their legal representative.

### Funding Information

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

### Conflict of Interest

Authors declares no conflicts of interest.

### References

- [1] Volkmann J, Mueller J, Deuschl G, Kühn A A, Krauss J K, Poewe W, Timmermann L, Falk D, Kupsch A, Kivi A, Schneider G-H, Schnitzler A, Südmeyer M, Voges J, Wolters A, Wittstock M, Müller J-U, Hering S, Eisner W, Vesper J, Prokop T, Pinsker M, Schrader C, Kloss M, Kiening K, Boetzel K, Mehrkens J, Skogseid I M, Ramm-Petersen J, Kemmler G, Bhatia K P, Vitek J L, Benecke R, and DBS study group for dystonia 2014 Pallidal neurostimulation in patients with medication-refractory cervical dystonia: a randomised, sham-controlled trial *Lancet Neurol* **13** 875–84
- [2] Meoni S, Fraix V, Castrioto A, Benabid A L, Seigneuret E, Vercueil L, Pollak P, Krack P, Chevrier E, Chabardes S and Moro E 2017 Pallidal deep brain stimulation for dystonia: a long term study *J Neurol Neurosurg Psychiatry* **88** 960–7
- [3] Vidailhet M, Vercueil L, Houeto J-L, Krystkowiak P, Benabid A-L, Cornu P, Lagrange C, Tézenas du Montcel S, Dormont D, Grand S, Blond S, Detante O, Pillon B, Arduin C, Agid Y,

- Destée A, Pollak P, and French Stimulation du Pallidum Interne dans la Dystonie (SPIDY) Study Group 2005 Bilateral deep-brain stimulation of the globus pallidus in primary generalized dystonia *N. Engl. J. Med.* **352** 459–67
- [4] Okun M S, Tagliati M, Pourfar M, Fernandez H H, Rodriguez R L, Alterman R L and Foote K D 2005 Management of referred deep brain stimulation failures: a retrospective analysis from 2 movement disorders centers *Arch Neurol* **62** 1250–5
- [5] Pauls K A M, Krauss J K, Kämpfer C E, Kühn A A, Schrader C, Stüdemeyer M, Allert N, Benecke R, Blahak C, Boller J K, Fink G R, Fogel W, Liebig T, El Majdoub F, Mahlknecht P, Kessler J, Mueller J, Voges J, Wittstock M, Wolters A, Maarouf M, Moro E, Volkmann J, Bhatia K P and Timmermann L 2017 Causes of failure of pallidal deep brain stimulation in cases with pre-operative diagnosis of isolated dystonia *Parkinsonism & Related Disorders* **43** 38–48
- [6] Reich M M, Horn A, Lange F, Roothans J, Paschen S, Runge J, Wodarg F, Pozzi N G, Witt K, Nickl R C, Soussand L, Ewert S, Maltese V, Wittstock M, Schneider G-H, Coenen V, Mahlknecht P, Poewe W, Eisner W, Helters A-K, Matthies C, Sturm V, Isaias I U, Krauss J K, Kühn A A, Deuschl G and Volkmann J 2019 Probabilistic mapping of the antidystonic effect of pallidal neurostimulation: a multicentre imaging study *Brain* **142** 1386–98
- [7] Vissani M, Isaias I U and Mazzoni A 2020 Deep brain stimulation: a review of the open neural engineering challenges *J. Neural Eng.*
- [8] Halpern C H, Danish S F, Baltuch G H and Jaggi J L 2008 Brain shift during deep brain stimulation surgery for Parkinson's disease *Stereotact Funct Neurosurg* **86** 37–43
- [9] Montgomery E B 2012 Microelectrode targeting of the subthalamic nucleus for deep brain stimulation surgery *Mov. Disord.* **27** 1387–91
- [10] Lozano C S, Ranjan M, Boutet A, Xu D S, Kucharczyk W, Fasano A and Lozano A M 2018 Imaging alone versus microelectrode recording-guided targeting of the STN in patients with Parkinson's disease *J. Neurosurg.* 1–6
- [11] Ben-Haim S, Asaad W F, Gale J T and Eskandar E N 2009 Risk factors for hemorrhage during microelectrode-guided deep brain stimulation and the introduction of an improved microelectrode design *Neurosurgery* **64** 754–62; discussion 762-763
- [12] Moran A, Bar-Gad I, Bergman H and Israel Z 2006 Real-time refinement of subthalamic nucleus targeting using Bayesian decision-making on the root mean square measure *Mov Disord* **21** 1425–31
- [13] Valsky D, Marmor-Levin O, Deffains M, Eitan R, Blackwell K, Bergman H and Israel Z 2017 Stop! Border Ahead: Automatic detection of subthalamic exit during deep brain stimulation surgery *Mov Disord* **32** 70–9
- [14] Vissani M, Cordella R, Micera S, Eleopra R, Romito L M and Mazzoni A 2019 Spatio-temporal structure of single neuron subthalamic activity identifies DBS target for anesthetized Tourette syndrome patients *J. Neural Eng.* **16** 066011
- [15] Micheli F, Vissani M, Pecchioli G, Terenzi F, Ramat S and Mazzoni A 2021 Impulsivity Markers in Parkinsonian Subthalamic Single-Unit Activity *Movement Disorders* **36** 1435–40
- [16] Koirala N, Serrano L, Paschen S, Falk D, Anwar A R, Kuravi P, Deuschl G, Groppa S and Muthuraman M 2020 Mapping of subthalamic nucleus using microelectrode recordings during deep brain stimulation *Scientific Reports* **10** 19241
- [17] Valsky D, Blackwell K T, Tamir I, Eitan R, Bergman H and Israel Z 2020 Real-time machine learning classification of pallidal borders during deep brain stimulation surgery *J Neural Eng* **17** 016021
- [18] Al Awadhi A, Tyrand R, Horn A, Kibleur A, Vincentini J, Zacharia A, Burkhard P R, Momjian S and Boëx C 2022 Electrophysiological confrontation of Lead-DBS-based electrode localizations in patients with Parkinson's disease undergoing deep brain stimulation *NeuroImage: Clinical* **34** 102971
- [19] Mosher C P, Mamelak A N, Malekmohammadi M, Pouratian N and Rutishauser U 2021 Distinct roles of dorsal and ventral subthalamic neurons in action selection and cancellation *Neuron* **109** 869-881.e6
- [20] Lu C W, Malaga K A, Chou K L, Chestek C A and Patil P G 2020 High density microelectrode recording predicts span of therapeutic tissue activation volumes in subthalamic deep brain stimulation for Parkinson disease *Brain Stimulation* **13** 412–9
- [21] Follett K A, Weaver F M, Stern M, Hur K, Harris C L, Luo P, Marks W J, Rothlind J, Sagher O, Moy C, Pahwa R, Burchiel K, Hogarth P, Lai E C, Duda J E, Holloway K, Samii A, Horn S, Bronstein J M, Stoner G, Starr P A, Simpson R, Baltuch G, De Salles A, Huang G D, Reda D J, and CSP 468 Study Group 2010 Pallidal versus subthalamic deep-brain stimulation for Parkinson's disease *N. Engl. J. Med.* **362** 2077–91
- [22] Lin S, Wu Y, Li H, Zhang C, Wang T, Pan Y, He L, Shen R, Deng Z, Sun B, Ding J and Li D 2020 Deep brain stimulation of the globus pallidus internus versus the subthalamic nucleus in isolated dystonia *Journal of Neurosurgery* **132** 721–32
- [23] Schiff S J, Dunagan B K and Worth R M 2002 Failure of single-unit neuronal activity to differentiate globus pallidus internus and externus in Parkinson disease *Journal of Neurosurgery* **97** 119–28
- [24] Starr P A, Turner R S, Rau G, Lindsey N, Heath S, Volz M, Ostrem J L and Marks W J 2006 Microelectrode-guided implantation of deep brain stimulators into the globus pallidus internus for dystonia: techniques, electrode locations, and outcomes *Journal of Neurosurgery* **104** 488–501
- [25] Alam M, Sanghera M K, Schwabe K, Lütjens G, Jin X, Song J, von Wrangel C, Stewart R M, Jankovic J, Grossman R G, Darbin O and Krauss J K 2016 Globus pallidus internus neuronal activity: a comparative study of linear and non-linear features in patients with dystonia or Parkinson's disease *J Neural Transm* **123** 231–40
- [26] Giorni A, Windels F, Stratton P G, Cook R, Silberstein P, Coyne T, Silburn P A and Sah P 2017 Single-unit activity of the anterior Globus pallidus internus in Tourette patients and posterior Globus pallidus internus in dystonic patients *Clin Neurophysiol* **128** 2510–8
- [27] Vitek J L, Chockkan V, Zhang J-Y, Kaneoke Y, Evatt M, DeLong M R, Triche S, Mewes K, Hashimoto T and Bakay R A E 1999 Neuronal activity in the basal ganglia in patients with generalized dystonia and hemiballismus *Ann Neurol* **46** 22–35
- [28] Tang J K H, Moro E, Mahant N, Hutchison W D, Lang A E, Lozano A M and Dostrovsky J O 2007 Neuronal Firing Rates

- and Patterns in the Globus Pallidus Internus of Patients With Cervical Dystonia Differ From Those With Parkinson's Disease *Journal of Neurophysiology* **98** 720–9
- [29] Cleary D R, Raslan A M, Rubin J E, Bahgat D, Viswanathan A, Heinricher M M and Burchiel K J 2013 Deep brain stimulation entrains local neuronal firing in human globus pallidus internus *J Neurophysiol* **109** 978–87
- [30] Starr P A, Rau G M, Davis V, Marks W J, Ostrem J L, Simmons D, Lindsey N and Turner R S 2005 Spontaneous Pallidal Neuronal Activity in Human Dystonia: Comparison With Parkinson's Disease and Normal Macaque *Journal of Neurophysiology* **93** 3165–76
- [31] Levi V, Franzini A, Rinaldo S, Coelli S, Bianchi A M, Franzini A, Nardocci N, Eleopra R and Zorzi G 2021 Globus pallidus internus activity during simultaneous bilateral microelectrode recordings in status dystonicus *Acta Neurochir* **163** 211–7
- [32] Neumann W-J, Horn A, Ewert S, Huebl J, Brücke C, Slentz C, Schneider G-H and Kühn A A 2017 A localized pallidal physiomaer in cervical dystonia *Ann. Neurol.* **82** 912–24
- [33] Piña-Fuentes D, van Zijl J C, van Dijk J M C, Little S, Tinkhauser G, Oterdoom D L M, Tijssen M A J and Beudel M 2019 The characteristics of pallidal low-frequency and beta bursts could help implementing adaptive brain stimulation in the parkinsonian and dystonic internal globus pallidus *Neurobiol. Dis.* **121** 47–57
- [34] Ewert S, Plettig P, Li N, Chakravarty M M, Collins D L, Herrington T M, Kühn A A and Horn A 2018 Toward defining deep brain stimulation targets in MNI space: A subcortical atlas based on multimodal MRI, histology and structural connectivity *NeuroImage* **170** 271–82
- [35] Erdman H B, Kornilov E, Kahana E, Zarchi O, Reiner J, Socher A, Strauss I, Firman S, Israel Z, Bergman H and Tamir I 2022 Asleep DBS under ketamine sedation: Proof of concept *Neurobiol Dis* **170** 105747
- [36] Lettieri C, Rinaldo S, Devigili G, Pualetto G, Verriello L, Budai R, Fadiga L, Oliynyk A, Mondani M, D'Auria S, Skrap M and Eleopra R 2012 Deep brain stimulation: Subthalamic nucleus electrophysiological activity in awake and anesthetized patients *Clin Neurophysiol* **123** 2406–13
- [37] Guang J, Baker H, Ben-Yishay Nizri O, Firman S, Werner-Reiss U, Kapuller V, Israel Z and Bergman H 2021 Toward asleep DBS: cortico-basal ganglia spectral and coherence activity during interleaved propofol/ketamine sedation mimics NREM/REM sleep activity *npj Parkinsons Dis.* **7** 67
- [38] Moran A and Bar-Gad I 2010 Revealing neuronal functional organization through the relation between multi-scale oscillatory extracellular signals *Journal of Neuroscience Methods* **186** 116–29
- [39] Quiroga R Q, Nadasdy Z and Ben-Shaul Y 2004 Unsupervised Spike Detection and Sorting with Wavelets and Superparamagnetic Clustering *Neural Computation* **16** 1661–87
- [40] Visser E, Keuken M C, Douaud G, Gaura V, Bachoud-Levi A-C, Remy P, Forstmann B U and Jenkinson M 2016 Automatic segmentation of the striatum and globus pallidus using MIST: Multimodal Image Segmentation Tool *NeuroImage* **125** 479–97
- [41] Solomon O, Palnitkar T, Patriat R, Braun H, Aman J, Park M C, Vitek J, Sapiro G and Harel N 2021 **Deep-learning based fully automatic segmentation of the globus pallidus interna and externa using ultra-high 7 Tesla MRI** *Hum Brain Mapp* **42** 2862–79
- [42] Leske S and Dalal S S 2019 Reducing power line noise in EEG and MEG data via spectrum interpolation *NeuroImage* **189** 763–76
- [43] Shannon C E 1948 A Mathematical Theory of Communication *Bell System Technical Journal* **27** 379–423
- [44] Panzeri S, Senatore R, Montemurro M A and Petersen R S 2007 Correcting for the Sampling Bias Problem in Spike Train Information Measures *Journal of Neurophysiology* **98** 1064–72
- [45] Magri C, Whittingstall K, Singh V, Logothetis N K and Panzeri S 2009 A toolbox for the fast information analysis of multiple-site LFP, EEG and spike train recordings *BMC Neurosci* **10** 81
- [46] Han J and Kamber M 2012 *Data mining: concepts and techniques* (Burlington, MA: Elsevier)
- [47] Breiman L 2001 Random Forests *Machine Learning* **45** 5–32
- [48] Quinlan J R 1986 Induction of decision trees *Mach Learn* **1** 81–106
- [49] Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Müller A, Nothman J, Louppe G, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M and Duchesnay É 2018 Scikit-learn: Machine Learning in Python *arXiv:1201.0490 [cs]*
- [50] Gholamy A, Kreinovich V and Kosheleva O 2018 Why 70/30 or 80/20 Relation Between Training and Testing Sets: A Pedagogical Explanation *International Journal of Intelligent Technologies and Applied Statistics* **11** 105–11
- [51] Rokai J, Rácz M, Fiáth R, Ulbert I and Márton G 2021 ELVISort: encoding latent variables for instant sorting, an artificial intelligence-based end-to-end solution *J Neural Eng* **18**
- [52] Radmanesh M, Rezaei A A, Jalili M, Hashemi A and Goudarzi M M 2022 Online spike sorting via deep contractive autoencoder *Neural Netw* **155** 39–49
- [53] Saif-Ur-Rehman M, Lienkämper R, Parpaley Y, Wellmer J, Liu C, Lee B, Kellis S, Andersen R, Iossifidis I, Glasmachers T and Klaes C 2019 SpikeDeeptector: a deep-learning based method for detection of neural spiking activity *J Neural Eng* **16** 056003