RESEARCH ARTICLE | APRIL 22 2025

Post-stroke spontaneous motor recovery in mice can be predicted from acute-phase local field potential using machine learning

Nicolò Meneghetti ; Michael Lassi ; Verediana Massa ; Silvestro Micera ; Alberto Mazzoni ; Claudia Alia ; Andrea Bandini

APL Bioeng. 9, 026108 (2025)
https://doi.org/10.1063/5.0263191

Post-stroke spontaneous motor recovery in mice can be predicted from acute-phase local field potential using machine learning

Cite as: APL Bioeng. **9**, 026108 (2025); doi: 10.1063/5.0263191
Submitted: 4 February 2025 · Accepted: 6 April 2025 ·
Published Online: 22 April 2025

Nicolò Meneghetti,[1,2,a)] Michael Lassi,[1,2] Verediana Massa,[3] Silvestro Micera,[1,2,4] Alberto Mazzoni,[1,2] Claudia Alia,[3] and Andrea Bandini[1,2,5]

AFFILIATIONS

[1]The Biorobotics Institute, Scuola Superiore Sant'Anna, Pisa, Italy
[2]Department of Excellence for Robotics and AI, Scuola Superiore Sant'Anna, Pisa, Italy
[3]CNR Neuroscience Institute, Pisa, Italy
[4]Modular Implantable Neuroprostheses (MINE) Laboratory, Università Vita-Salute San Raffaele & Scuola Superiore Sant'Anna, Milan, Italy
[5]Health Science Interdisciplinary Research Center, Scuola Superiore, Sant'Anna, Pisa, Italy

[a)]**Author to whom correspondence should be addressed:** nicolo.meneghetti@santannapisa.it

ABSTRACT

Stroke remains a leading cause of long-term disability, underscoring the urgent need for effective predictors of motor recovery. Understanding the electrophysiological changes underlying spontaneous recovery could offer critical insight into recovery mechanisms and aid in predicting individual rehabilitation trajectories. In this study, we investigated the predictive power of local field potentials recorded 2 days post-stroke to forecast 1 month motor recovery in a mouse model of ischemic stroke. By employing a comprehensive machine learning approach, we identified key electrophysiological features that significantly enhanced prediction accuracy. Through nested leave-one-animal-out cross-validation, we achieved high prediction accuracy, correctly identifying motor recovery status in 15 out of 16 mice. Our findings also revealed that pre-stroke brain activity did not contribute to prediction accuracy, suggesting that post-stroke dynamics are the primary determinants of recovery. Notably, we found that features from the contralesional hemisphere were particularly influential in predicting recovery outcomes, underscoring the critical role of the non-lesioned hemisphere in motor recovery. Our data-driven methodology underscores the importance of balancing feature selection to optimize predictive performance, particularly in the context of spontaneous recovery, where insight into natural recovery processes can guide the development of targeted rehabilitation strategies. Ultimately, our findings advocate for a deeper understanding of post-stroke brain dynamics to improve clinical outcomes for stroke patients.

I. INTRODUCTION

Stroke is one of the leading causes of disabilities worldwide.[1] Most stroke survivors suffer from disabilities chronically after stroke, especially in the motor domain.[2] For this reason, stroke poses a huge economic burden on healthcare systems, due to its treatment and rehabilitation costs.[3] Hence, predicting the potential for recovery after stroke is essential to guide the best rehabilitation path, improve the quality of life of the patients, and minimize societal impact.[3]

Stroke neuronal loss leads to changes in the plasticity processes immediately after the event, both locally in the lesioned area and globally in terms of rearrangement of synaptic connections to further brain areas.[4,5] In fact, the stroke lesion can trigger the opening of a window of enhanced plasticity,[6] in which partial spontaneous recovery occurs from the acute to the chronic stroke phases.[7,8] However, not all stroke survivors display the same degree of post-stroke recovery of motor deficits. Several factors influence motor rehabilitation outcome, such as lesion volume and location,[9] age at stroke onset,[10,11] and initial severity of motor deficits.[12] Hence, knowing in advance the patient-specific potential for recovery is of utmost importance to optimize the rehabilitation plan as early as possible after the stroke. For instance,

knowledge of a patient's recovery prognosis could help in crafting rehabilitation objectives that are in line with the patient's potential for independence or need for assistance, thereby conserving resources and maximizing healthcare outcomes.

The simplest stroke recovery model, known as the proportional recovery rule, solely relies on the initial impairment of the patients and state that the amount of recovery at the end of the subacute stage after stroke is proportional to the level of initial impairment.[13,14] Specifically, it was observed that most stroke patients recover about 70% of their total initial motor impairment, as measured by clinical scales, such as the Fugl–Meyer Assessment Scale. However, not all patients follow this recovery rule, especially when the degree of their initial impairment is high.[15] The proportional model has also faced criticism due to the ceiling effects observed in clinical scales.[16] These scales tend to saturate toward their maximum values at lower levels of impairment, leading to a mathematical coupling between initial impairment and subsequent recovery.[16] This coupling partially explains the recovery rule, raising questions about the model's validity, limiting the scales' ability to predict actual patients' motor recovery. At the same time, more complex methods to estimate stroke recovery in humans often require evaluating the integrity of the corticospinal tract,[17] along with the clinical assessment of residual motor function.[18] This evaluation is performed with either structural brain imaging[19] or transcranial magnetic stimulation,[20] which are expensive and not readily available in all healthcare centers. These considerations highlight the need for a widely available and reliable tool to predict stroke recovery from neurophysiological recordings.

In this direction, the use of electrophysiological recordings may be a cost-effective way to monitor stroke motor recovery, offering a window on the functional potential for recovery of the patient. Moreover, given the portability of many electrophysiological recording systems, these potential biomarkers may then be used to guide and adapt treatments in a patient-tailored manner.[21,22] In humans, there is increasing evidence that noninvasive functional brain signals, such as electroencephalogram (EEG), can be used to track recovery in stroke patients.[23–25] Specifically, resting-state measures of low-to-high frequency ratios in the EEG activity (such as the delta-to-alpha ratio and delta/theta-to-alpha/beta ratio) have been shown to be predictive of the recovery from neurological deficits.[26–28] Measures of spectral brain symmetry, such as the pairwise derived brain symmetry index and its variants are predictive of recovery from motor deficits after stroke.[29–31] Power spectrum,[32] interhemispheric connectivity,[33–35] network metrics,[36] and signal complexity[37] also change throughout stroke recovery phases and have a varying degree of predictive ability of various aspects of the recovery process.[23]

While human studies have revealed the potential of using electrophysiological recordings to predict motor recovery after stroke, investigating the "true" spontaneous recovery in humans is challenging due to ethical considerations. Standard clinical practice requires patients to undergo rehabilitation early after stroke, which is essential for patient care but may act as a confounding factor when studying spontaneous recovery. Furthermore, pre-stroke electrophysiological data for human patients is typically unavailable. Consequently, mouse models are invaluable for examining "true" spontaneous recovery and evaluating the predictive value of pre-stroke factors in recovery processes. While EEG is a common electrophysiological technique in humans, invasive recording methods are more accessible in mice. Specifically, local field potentials (LFPs) are frequently utilized to gain insight into the activity of specific brain areas in mice, potentially offering a more detailed understanding of the electrophysiological changes associated with stroke and recovery.

LFPs reflect transmembrane current flow in neuronal populations near the electrode tip, typically within hundreds of micrometers.[38] These signals are thought to primarily represent synaptic inputs to the tissue surrounding the electrode,[39] providing a localized window of observation of neural activity. While LFPs offer less information than EEG about widespread cortical dynamics, they are considered one of the most informative signals for studying cortical electrogenesis.[38] Given their specificity and sensitivity to local neural activity, LFPs represent a powerful tool for identifying relevant biomarkers linked to a wide range of pathological conditions,[39] including stroke. In mouse stroke models, LFP recordings can uncover robust electrophysiological markers that facilitate the study of spontaneous recovery. Specifically, this would allow to effectively study the processes of spontaneous recovery after a stroke, as no treatment needs to be applied to the mouse. Moreover, stroke mouse models would enable to more effectively distinguish between true spontaneous recovery,[40] compensatory behavioral mechanisms[41] and rehabilitation-induced functional brain reorganization.[8] In addition to that, pre-stroke activity may be studied in mice models to highlight protective factors that are influential in the recovery processes. LFP activity can be studied in freely behaving mice, a self-determined behavior that loosely matches the resting-state recordings in humans.[42]

In mice studies, the reorganization of ipsi- and contra-lateral activity was observed in LFP, specifically with a reduction of gamma band activity in the peri-lesional area, with a related increase in power in the same band in the contra-lateral hemisphere.[43] At the same time, an increased low-frequency oscillatory activity can be used to track stroke recovery as well.[22] Another study identified LFP complexity, as measured by the entropy measure, as a target biomarker for tracking recovery.[44] Abnormal imbalance of inter-hemispheric interaction was also observed in the sub-acute to chronic stages after stroke when compared to non-lesioned mice.[45]

However, to the best of our knowledge, both in humans and in mice models, only a few studies evaluated the ability of electrophysiological biomarkers to predict spontaneous future recovery in a completely data-driven manner.[23,46–48] Therefore, our work aimed at studying the electrophysiological correlates of forelimb motor recovery in a mouse model of middle cerebral artery occlusion (MCAO), main cause of ischemic insult in humans, by extracting a set of features pertaining to spectral, interhemispheric connectivity, and signal complexity domains. We aimed at developing a machine learning model to forecast motor recovery, by combining the most predictive univariate features into a multivariate classifier, able to distinguish good and poor recoverers. Specifically, our goal was to verify which features in the ipsi- and contra-lesional local field potentials activity and in their interaction are sufficient to predict motor recovery. Finally, we aimed at investigating whether pre-stroke baseline activity carries information about future recovery.

II. RESULTS

A. Two-day post-stroke LFPs accurately forecast 1-month motor recovery

To investigate the role of acute changes in electrophysiological activity in the perilesional and contralesional motor cortices as

predictors of spontaneous post-stroke recovery, we employed the following experimental design [Fig. 1(a)]: 16 mice underwent implantation of two chronic bipolar electrodes in the caudal forelimb areas (CFA) of both hemispheres. LFP recordings were obtained under freely moving conditions, alongside motor performance assessments using the Motor Score (see the methods), at baseline (i.e., pre-stroke), and at 2 and 30 days following permanent middle cerebral artery occlusion (MCAO) induction. The Motor Score obtained at 30 days, compared to baseline values, was used to classify the animals into good and poor recoverers.

To evaluate the predictability of motor recovery from extracellular recordings, we analyzed the LFP recorded 2 days post-stroke in freely behaving mice. Following artifact rejection due to animal movements (see Sec. IV), the LFPs were segmented into windows, and a set of 28 features was extracted from each window [Fig. 1(b)]. When visualized in a lower-dimensional space [two dimensions using t-SNE,[49]

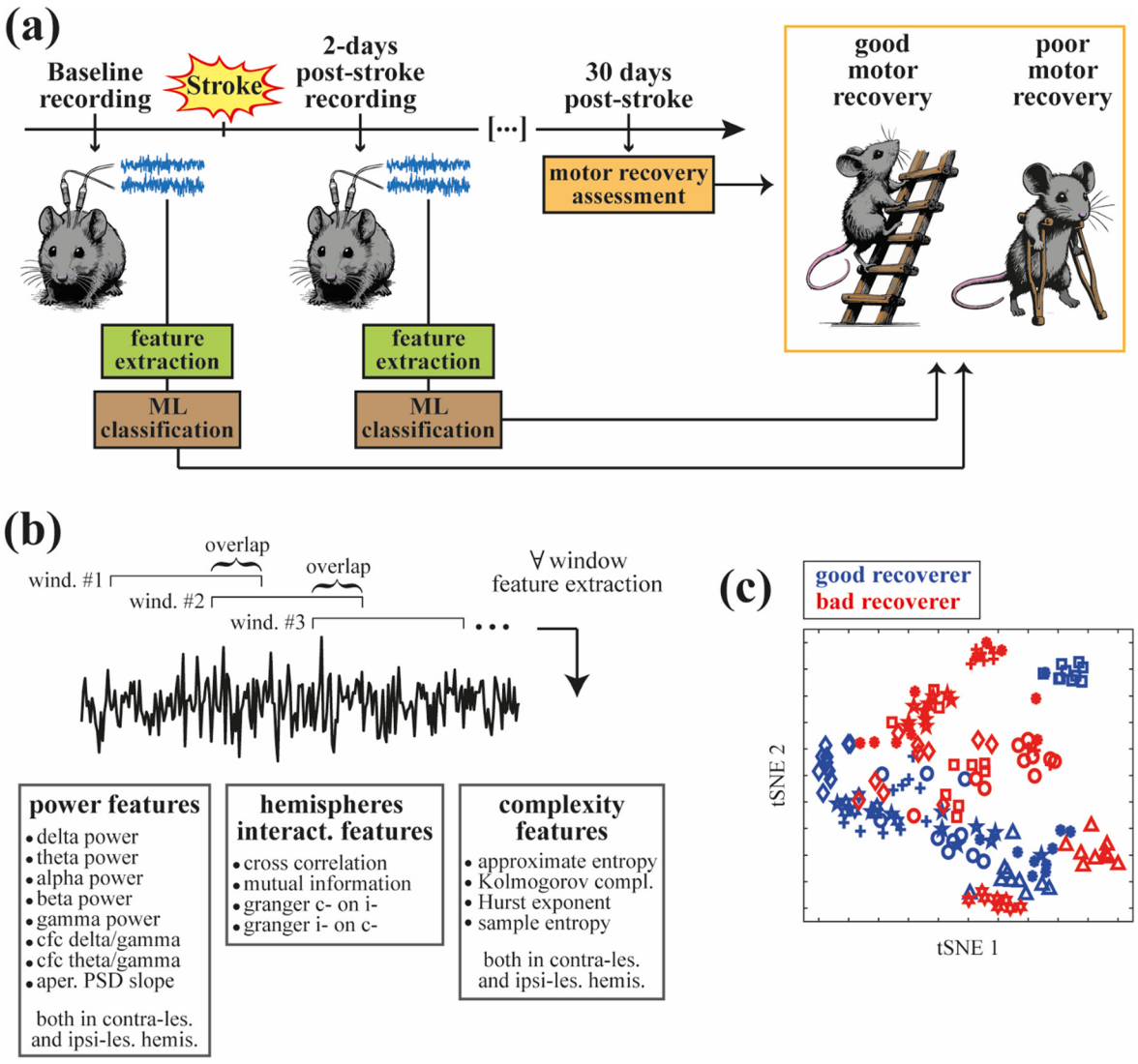

FIG. 1. Experimental paradigm and feature extraction. (a) Schematic of the experimental paradigm (see Sec. IV). Sixteen mice (N = 16) were recorded using two electrodes implanted in the motor cortex of both hemispheres prior to permanent middle cerebral artery occlusion. A set of features was extracted from the collected signals and fed into a machine learning classification algorithm. The mice underwent a battery of motor tests both before and 30 days after the induced stroke. A comparison of cumulative motor performance was then used to sort the mice into good and poor recovery groups. (b) Feature extraction from the collected extracellular signals was based on a sliding window approach (see Sec. IV) of varying size (30 or 60 s) and degree of overlap (50%, none, or random). For each temporal window we extracted a set of 28 features grouped into three categories: 8 based on spectral power (for each hemisphere); 4 accessing inter-hemispheric interaction; and 4 based on signal complexity (for each hemisphere). (c) t-SNE plot of the dataset obtained 2 days post-stroke. Each marker represents a 60-s window with no overlap. Markers of the same shape indicate windows from the same animal. Red markers denote poor recoverers, while blue markers indicate good recoverers.

see Fig. 1(c)], the features obtained 2 days post-stroke appear to form two distinct clusters, corresponding to poor and good recoverers, with minimal overlap between clusters. This observation provided an initial indication that data-driven approaches could potentially be used to infer recovery class.

A nested leave-one-animal-out cross-validation (LOAO-CV) procedure was employed to optimize the multiple hyperparameters adopted for classification. We iteratively excluded one animal at a time to identify the optimal window lengths (30 and 60 s), overlap strategies (no overlap, 50% overlap, and random overlap), and classification algorithms [support vector machine (SVM) with four different kernel functions, random forest, artificial neural network (ANN), and k-nearest neighbor (kNN)] along with their respective hyperparameters (see Sec. IV). The best hyperparameters were those that maximized the animal-wise classification accuracy on the validation set using a LOAO-CV strategy (see Sec. IV). Subsequently, the model with the best-performing hyperparameters was tested on the features of the excluded animal. This process was repeated for each of the 16 animals.

Across the 16 mice, the best-performing classifier was consistently identified as the SVM. Specifically, the linear kernel was adopted in 11 of the 16 mice, while the remaining 5 utilized the radial basis function kernel. The optimal window length was consistently 60 s for all mice, with a no-overlap strategy preferred in the majority of cases (14 out of 16 mice). Regarding the selected hyperparameters, the C parameter was set to 10 in 6 animals, 100 in 8 animals, and 1000 in 2 animals. The gamma parameter was set to 10–3 in 15 animals and to 10 in 1 animal. The average LOAO-CV accuracy on the validation set was [82.5 ± 1.9]% of correctly classified animals (mean ± standard error of mean, sem), while [78.7 ± 1.5]% (mean ± sem) of correctly classified windows.

Finally, we tested the classifier (with the previously identified best-identified hyperparameters) on the animal excluded from each validation step. We found that the features collected from the windowed LFPs recorded 2 days post-stroke correctly classified the 30-day post-stroke motor recovery in 13 out of 16 animals [accuracy: 81.3%, precision: 85.7%, recall: 75.0%, F1-score: 80.0%, Fig. 2 point (3)]. The accuracy of the classification based on the windows was 74.6% (precision: 77.0%, recall: 70.0%, and F1-score: 73.3%).

To ensure that our prediction results were not dependent on the cross-validation strategy used, we performed the classification task using a conventional fixed-split cross-validation scheme (fixed-CV, see Methods). We found that the prediction accuracy was comparable to the previous leave-one-animal-out approach, correctly identifying 5 out of 6 animals of the test set, and 75.0% of windows (Fig. S1).

B. Tradeoff between feature dimensionality and prediction performance

We sought to determine the minimal set of features required to uphold classification accuracy.

To assess the relative importance of each feature in predicting 1-month motor recovery, we adopted the maximum relevance minimum redundancy (mRMR) algorithm to perform a leave-one-animal-out (LOAO) backward feature selection. Briefly, for each animal excluded from the dataset, we performed an mRMR feature ranking on the dataset obtained from the remaining 15 animals. This process was repeated for each animal in the dataset.

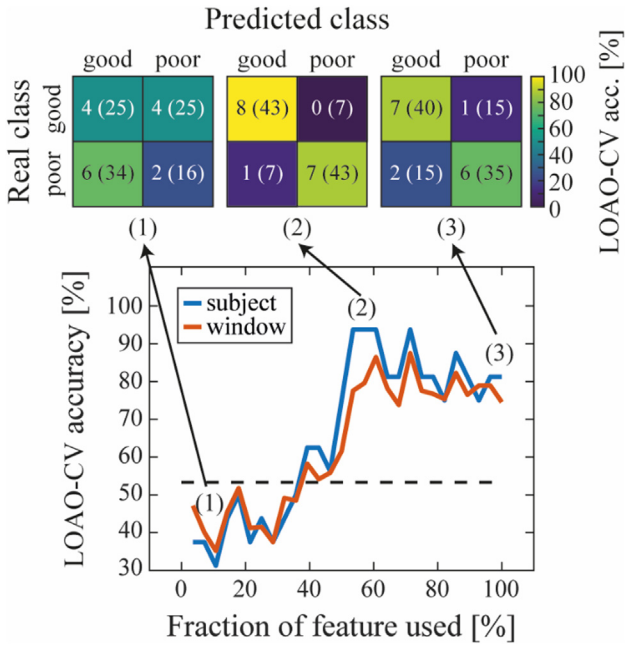

FIG. 2. Accuracy of motor recovery prediction at 2 days post-stroke. *Top*: Confusion matrices for the leave-one-animal-out cross-validation (LOAO-CV) in three different fractions of features adopted: using only the most important feature [i.e., point (1)], the half most important features [i.e., point (2)], and all 28 features [i.e., point (3)]. The numbers outside parenthesis within the confusion matrices entries indicate the classified animals, while the numbers within parentheses represent the classified windows. *Bottom*: Accuracy of motor recovery using LOAO-CV across different fractions of features employed for classification. Features were incrementally added in order of their importance, starting from the most important [i.e., point (1)] to the least important, until all 28 features in the dataset were utilized [point (3)].

Similar to the procedure described in Sec. II A, we first conducted a LOAO-CV step. Specifically, by excluding one animal at a time, we identified the set of hyperparameters that maximized animal-wise classification accuracy. This time, the process involved iteratively discarding the least important feature at each step, as determined by the mRMR algorithm, until only the most important feature remained for the classification. Finally, the classifier with the best hyperparameters for each combination of feature set size were tested on the excluded animal.

Interestingly, we found that throughout all the LOAO iterations and feature combinations employed, the optimal classifier was the SVM in 97.1% of cases (44.4% using the linear kernel, 42.5% using the radial basis function, 7.8% using the quadratic kernel, and 5.3% using the cubic kernel). Regarding the optimal window, the 60-s window was preferred 77.9% of the time, with a no-overlap strategy being optimal in 82.4% of cases.

We found that the LOAO test classification accuracy remained at chance level [Fig. 2, point (1)] until approximately half of the most important features were retained for the classification [Fig. 2, point (2)]. Accordingly, the average accuracy for the first half of the most important features [i.e., from point 1 to point 2 in Fig. 2] was [47.8 ± 3.3]% (mean ± sem). When 15 (out of 28) of the most important features were employed, the classification accuracy reached a

maximum, correctly classifying 15 out of 16 animals [Fig. 2, point (2)]. Including the remaining less relevant features slightly decreased the testing accuracy, achieving a level of 13 out of 16 correctly classified animals [Fig. 2, point (3)], which corresponds to the classification result illustrated in Sec. II A (where all the features were employed).

A similar trend was observed with the accuracy based on windows. Until half the features were adopted, the window-based accuracy remained at chance level, with an average of [47.7 ± 2.1]% (mean ± sem) for the first half of the most important features. The window-based accuracy reached its peak of 86.5% when 17 (out of 28) most important features were employed [Fig. 2, point (2)] and showed a slight decrease when all the features were used, achieving 74.6% accuracy [Fig. 2, point (3)].

A similar trend was observed when performing the classification task using a conventional cross-validation scheme. The CV subject-wise accuracy remained at chance level, averaging [53.6 ± 2.6]% (mean ± sem) for the first half of the most important features [Fig. S1, point (1)]. The subject accuracy reached its maximum of 6 out of 6 correctly classified test-set animals when at least the most relevant 14 features were used [Fig. S1, point (2)], while slightly decreasing when adding the remaining (approximately) half of less relevant features [achieving 5 out of 6 correctly classified animals in the test set, Fig. S1, point (3)]. Comparable results were obtained with the accuracy based on windows (red trace in Fig. S1).

C. The role of contralesional LFPs and data-driven approaches in predicting post-stroke motor recovery

To validate the feature importance identified through our machine learning approach, we compared their relevance using traditional statistical methods.

Each LOAO iteration (where each animal sequentially served as the test set) produced its own feature ranking. To derive an overall feature ranking, we then averaged the rankings of each feature as determined by the mRMR algorithm across all animals [Fig. 3(a)]. In parallel, we performed statistical tests to determine if the importance of features identified by the machine learning approach was also reflected in first-order statistical tests across the two classes of motor recovery. To this end, we conducted pairwise comparisons for each feature between good and poor motor recovery mice using a two-sided Wilcoxon rank-sum (WRS) test [Fig. 3(a), green trace]. Multiple comparisons were corrected for using the Bonferroni approach. Given that the optimal window strategy identified in every LOAO iteration during validation was 60 s long with no overlap for most animals, we pooled feature values across this window strategy.

We found a significant correlation (Pearson $\rho = -0.71$, $p < 0.001$) between the p-values of the Wilcoxon rank-sum (WRS) test across features and the average ranking returned by the mRMR algorithm.

Interestingly, only a very small subset of features displayed statistical significance between poor and good recoverers [asterisks in Fig. 3(a)]. Presented in descending order of their average ranking, the Wilcoxon rank-sum test identified the following five features as significantly different [Fig. 3(b)]: (i) the Granger causality test-statistics of contralateral LFPs on ipsilesional LFPs; (ii) approximate entropy of contralesional LFPs; contralesional (iii) alpha, (iv) theta, and (v) gamma power.

This set of five significant features [Fig. 3(b)] highlights the importance of the electrophysiological properties of the contralesional hemisphere. Four out of the five significant features were exclusively related to contralesional LFPs, and the last reflected the impact of contralesional LFP on the perilesional hemisphere. The higher importance of contralesional LFP features was further supported by their higher average ranking when considering the overall set of 28 features [Fig. 3(a): contralesional average rank—orange bars: 13.8 ± 2.0, median rank ± SEM; ipsilesional average rank—gray bars: 21.0 ± 1.8, median rank ± SEM].

In total, 28 features were extracted from the recorded LFPs, categorized into power, complexity, and interhemispheric interaction [Fig. 1(b)]. The five statistically significant features spanned all three categories, indicating no specific preference for any one category: three were power-based, one reflected interhemispheric interaction, and one pertained to signal complexity. In line with this, the average ranking of all 28 features within each category did not show a clear preference either, with no significant difference between the categories [Fig. 3(a): power features average rank: 16.1 ± 1.9; complexity features average rank: 13.2 ± 2.4; interhemispheric communication features average rank: 17.7 ± 1.3, mean rank ± SEM].

Up to this point, we have focused on feature values extracted from LFP recordings obtained 2 days post-stroke. We next examined whether these features also exhibited significant changes relative to baseline (pre-stroke) values [Fig. 3(c)]. A mixed-design ANOVA was employed to evaluate the effects of two factors: stroke (baseline vs 2 days post-stroke, i.e., the within subject factor) and group (good vs poor recoverers, i.e., the between subject factor), along with their interaction.

We found that the stroke factor had a significant effect on the four contralesional features, indicating that these features changed significantly following the stroke. Specifically, approximate entropy ($p < 0.001$), alpha power ($p < 0.001$), theta power ($p < 0.001$), and gamma power ($p = 0.002$) showed significant differences between baseline and post-stroke conditions. However, the group factor also had a significant effect on all five features, with the exception of theta power, where no significant difference was observed. The features that were significantly affected by group included Granger causality ($p < 0.001$), approximate entropy ($p = 0.006$), alpha power ($p = 0.04$), and gamma power ($p < 0.001$).

Notably, the interaction between stroke and group was significant for four of the five features, indicating that the effect of stroke on these features differed between the good and poor recoverers. This suggests that the temporal evolution of these features was influenced not only by the presence of stroke but also by the recovery status of the animals. Specifically, the interaction effects were significant for Granger causality ($p < 0.001$), approximate entropy ($p < 0.001$), alpha power ($p < 0.001$), and theta power ($p < 0.001$), highlighting differential patterns of feature changes between the two groups over time.

To further investigate this aspect, post-hoc tests were performed to determine whether there were significant differences between pre- and post-stroke values for each group. For mice with good motor recovery, no statistically significant differences were found between pre- and post-stroke values for any of the features [Fig. 3(c)]. In contrast, mice with poor motor recovery showed statistically significant differences between pre- and post-stroke values [Fig. 3(c)] for the four most important features (i.e., Granger causality, approximate entropy, alpha, and theta power). This indicates that the good recovery group

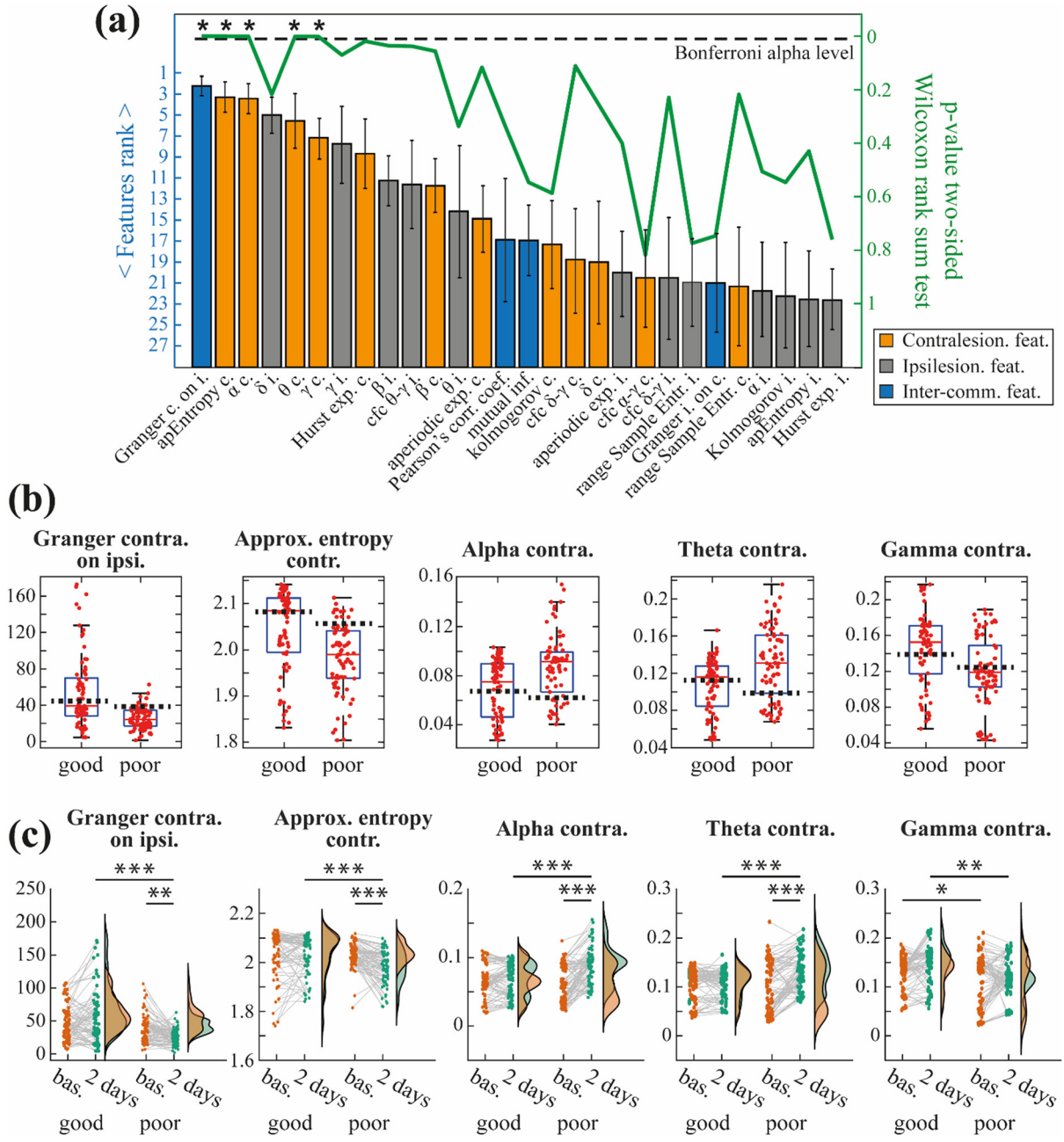

FIG. 3. Feature importance at 2 days post-stroke for predicting motor recovery. (a) (Left y-axis) Average rank across animals computed with the *mrmr* algorithm. The feature names are indicated on the x-axis (description of features' abbreviations can be found in the Sec. IV). Error bars indicate standard deviation of feature ranking. Bars were color-coded according to the feature category: orange for contralesional features, gray for ipsilesional feature and blue for inter-hemispheric communication. (right y-axis) P-values (green line) for the pairwise comparison between good and poor motor recovery for each feature, computed with the two-sided Wilcoxon rank sum test. Multiple comparisons were accounted for by Bonferroni method. The dashed line indicates alpha value corrected with the Bonferroni method. Asterisks indicate statistical significance. (b) Scattered boxplots for the five statistically significant features between good and poor motor recoverers. From left to right: (i) the Granger causality test-statistics of contralateral LFPs on ipsi-lesional LFPs (Wilcoxon rank-sum test $p < 0.001$); (ii) approximate entropy of contra-lesional LFPs (Wilcoxon rank-sum test $p < 0.001$); (iii) alpha contra-lesional power (Wilcoxon rank-sum test $p = p < 0.001$); (iv) theta contra-lesional power (Wilcoxon rank-sum test $p < 0.001$); and (v) gamma contra-lesional power (Wilcoxon rank-sum test $p < 0.001$). Each point represents a different window. Dashed black line represent the mean value of the corresponding motor recovery group in baseline. Error bars represent standard deviation. (c) Distribution of the five statistically significant features between good (left) and poor (right) motor recoverers at 2 days post stroke ("2 days," green) and before the stroke ("bas." short for baseline, orange). Each point represents the feature value within a single time window. Gray lines connecting the points indicate paired observations, linking pre- and post-stroke values. Probability density functions are reported vertically. Asterisks indicate the significance of Tukey's honest significant difference post-hoc tests: $^{*}p < 0.05$, $^{**}p < 0.01$, and $^{***}p < 0.001$.

maintained stability in these features after the stroke, whereas the poor recovery group experienced changes, suggesting that stroke has a differential impact on feature evolution depending on recovery outcomes.

In the previous analysis, feature values were extracted from multiple 60-s, non-overlapping windows, providing several data points per animal. To assess the impact of averaging these values across animals, we conducted additional analyses using animal-averaged feature values for each time window (Fig. S2). We first performed a Wilcoxon rank-sum pairwise comparison on these averaged values, which revealed no significant differences between the poor and good motor recovery groups post-stroke. This suggests that averaging the feature values across animals may obscure meaningful group differences. Moreover, we applied the same mixed-design ANOVA to the averaged dataset to access differences between good and poor motor recoverers and between pre- and post-stroke condition. We found that only Granger causality was significantly modulated by the group factor ($p = 0.04$), while the other features did not show statistical significance when averaged across animals. These findings underscore the importance of retaining multiple data points per animal by using the windowed approach to enhance data numerosity and statistical power.

D. Pre-stroke LFPs fail to predict motor recovery at 1 month

So far, we have focused on forecasting 1-month post-stroke outcomes based on features extracted from LFPs recorded 2 days after the stroke. We also investigated whether baseline recordings (i.e., LFPs collected before the stroke lesion) held any predictive value for motor recovery.

Following a similar methodology as for the 2-day post-stroke recordings, we performed a nested LOAO-CV to optimize multiple hyperparameters for the baseline LFPs. Consistently with the previous results, the SVM emerged as the best classifier for all 16 animals. However, the optimal window length and overlap strategies were more variable compared to the 2-day post-stroke scenario: in 12 out of 16 animals the 60-s window was the optimal one, while the overlap strategies were diverse (no overlap: 6 animals, 50% overlap: 5 animals, and random overlap: 5 animals).

This heterogeneity in hyperparameters was reflected in a significant decrease in validation performance: the animal-wise accuracy was on the validation set was of $[61.7 \pm 2.2]\%$ (mean $\pm$ sem), while the window-based accuracy was $[56.3 \pm 1.8]\%$.

Next, we iteratively tested the best-identified hyperparameters on the excluded animal of each LOAO validation step. We found that the pre-stroke LFPs failed to correctly classify motor recovery classes in almost every animal, with performance falling significantly below the chance level (i.e., 50%). Specifically, only 2 out of 16 animals (animal-wise accuracy: 12.5%) were correctly classified. Similarly, only 11.9% of the windows were correctly classified using baseline LFP features.

Interestingly, randomizing the feature values (while maintaining the same labeling patterns) yielded similar results. Correct classification was achieved in only 3 out of 16 animals (animal-wise accuracy: 18.8%) and 26.0% of windows with the random dataset.

These results, falling significantly below the expected 50% chance level in balanced two-class classification tasks, can be attributed to the interaction between the LOAO testing procedure and the label assignment based on majority voting from windows. Briefly, in each LOAO iteration, the training set was obtained by excluding one animal at a time, resulting in an unbalanced training set biased toward the class opposite to that of the test animal. The classification was then performed on the windows of the test animal, with the final label assigned based on the majority vote from its windows, which was disproportionate in favor of the other animal class (see Materials and Methods for a detailed description). This resulted in a systematic bias toward the class opposite to that of the test animal, pushing the test classification accuracy significantly below the chance level.

To address this issue stemming from the imbalance created by the LOAO-CV procedure, we performed a leave-two-animals-out cross-validation (L2AO-CV) procedure. Briefly, every L2AO iteration consisted of iteratively removing from the validation/training set a couple of randomly picked animals of opposite motor recovery class (see Sec. IV).

Following this new validation scheme, we first performed validation and testing on the 2 days post-stroke dataset. The average L2AO testing accuracy across animals for the 2 days post-stroke recordings was $[81.3 \pm 4.2]\%$ correctly classified animals (precision: $[80.0 \pm 6.5]\%$, recall: $[75.0 \pm 7.6]\%$; F1-score: $[80.0 \pm 5.3]\%$), in line with the results found using the LOAO procedure (see Sec. III B). For the pre-stroke LFPs, instead, the average L2AO accuracy across animals reached $[41.6 \pm 3.5]\%$ (precision: $[43.7 \pm 4.0]\%$, recall: $[56.3 \pm 6.9]\%$; F1-score: $[52.6 \pm 3.8]\%$). This result highlights (1) the lack of informative power regarding motor recovery in the LFP features extracted before the stroke and (2) the existence of a bias affecting LOAO accuracy when approaching the condition of no difference between the two animal recovery groups.

Finally, we tested whether normalizing the features extracted at 2 days post-stroke for each animal by their corresponding pre-stroke values could improve classification accuracy. With the LOAO-CV scheme, the classification accuracy dropped to 8 out of 16 correctly classified animals. Using the L2AO, the average accuracy across subjects reached $[65.7 \pm 4.0]\%$ (precision: $[68.5 \pm 6.7]\%$, recall: $[59.3 \pm 7.6]\%$; F1-score: $[66.7 \pm 4.3]\%$), which is intermediate between the accuracy obtained using the 2-days post-stroke and pre-stroke LFP features.

III. DISCUSSION

Our study demonstrated that LFPs recorded 2 days post-stroke can predict 1-month spontaneous motor recovery in mice with excellent classification performance (Fig. 2). By employing a comprehensive machine learning approach (Fig. 1), we identified key electrophysiological features that significantly contribute to this prediction (Fig. 3). Notably, our findings underscore the critical role of contra-lesional brain activity, and the efficacy of data-driven methodologies in understanding and forecasting post-stroke spontaneous motor recovery. Interestingly, we found that incorporating pre-stroke brain activity did not improve the classification accuracy. This indicates that motor recovery is mainly determined by the stroke aftermath and suggests the absence of a preexisting neuroprotective effect.

A. Predictive power of post-stroke LFP

One of the main findings of our study is the significant predictive power on spontaneous motor recovery of acute post-stroke local field potentials. Our machine learning algorithms, trained on features extracted from LFPs recorded 2 days post-stroke, demonstrated remarkable accuracy in predicting motor recovery outcomes 1 month

after the lesion. Such a prediction achieved a high classification rate, correctly predicting recovery status in 15 out of 16 animals using a nested leave-one-animal-out paradigm [Fig. 2]. This performance underscores the potential of early post-stroke LFPs as a robust predictor of long-term recovery.

Our machine learning approach utilized a method of windowing the electrophysiological signal estimating features within each window [Fig. 1(b)]. We tested two window lengths (30 and 60 s) and various overlap configurations (50%, 0%, and random overlap) between neighboring windows. This design allowed to increase the number of data points compared to estimating a single feature value for each tested animal.

Using a nested leave-one-animal-out cross-validation paradigm, we identified the optimal window length and overlap for each animal. Interestingly, the optimal configuration for each tested animal was the longest window length (60 s) with no overlap between neighboring windows. This result highlights the necessity of balancing two factors: increasing the dataset size, ensuring reliable feature estimation, and managing feature correlation.

Maximizing the dataset size by using the shortest window length would have provided the highest number of data points, and similarly, utilizing maximum overlap would have further increased the data volume. However, these strategies could compromise the reliability of feature estimation and introduce higher correlations between features, which might adversely affect prediction performance. The selection of a 60-s window without overlap strikes an effective balance, yielding enough data for robust feature estimation while minimizing the risk of excessive feature correlation. This balance is crucial for extracting meaningful features that enhance the accuracy of our machine learning model in predicting motor recovery outcomes.

Finally, we investigated the relative importance of the features used in predicting motor recovery. Our analysis revealed that achieving high prediction accuracy required employing the most informative half of the features (Fig. 2). Including the less informative half did not improve prediction accuracy; in fact, it showed a slight trend toward reduced accuracy. This finding underscores the value of a data-driven approach that prioritizes a carefully selected, balanced set of features over reliance on single biomarkers.

B. Stroke-induced changes vary between different recovery trajectories

As previously noted, one of the main findings of our approach is the absence of single biomarkers for spontaneous motor recovery. Accurate prediction of motor recovery was only achieved through a data-driven approach that combined several features.

Nevertheless, we identified a limited subset of features that consistently emerged as the most important across different mice [Fig. 3(a)]. Among these, five features were statistically significant in distinguishing between animals with poor and good motor recovery at 2 days post-stroke: (i) the Granger causality test-statistics of contralesional on ipsilesional LFPs; (ii) approximate entropy of contralesional LFP; (iii) contralesional alpha, (iv) theta, and (v) gamma LFP power.

Interestingly, by incorporating pre-stroke baseline information, we identified that the primary distinction between good and poor motor recoverers lay in the stability of such features following stroke [Fig. 3(c)]. Specifically, good recoverers were characterized by the preservation of these features, which remained largely unaffected by stroke. In contrast, poor recoverers exhibited marked deviations from their baseline levels, including a decrease in interhemispheric Granger causality and entropy, alongside an increase in theta and alpha power oscillations [Figs. 3(b) and 3(c)].

When analyzing subject-averaged data for these features, none of the pairwise statistical comparisons between good and poor recoverers at 2 days post-stroke rejected the null hypothesis, suggesting no significant differences at this stage. This highlights the critical value of retaining multiple data points per animal through the windowed analysis approach, which increases data numerosity and enhances statistical power. Among the features examined, only the most significant (i.e., Granger causality test statistics of contralesional on ipsilesional LFP) demonstrated a robust group effect (mixed-design ANOVA, $p = 0.01$) in the subject-averaged dataset, underscoring its potential as a distinguishing factor between recovery trajectories. This finding supports the idea that changes in inter-hemispheric functional coupling immediately following a stroke are key biomarkers of both stroke-induced deficits, as previously found by Hazime *et al.*,[43] as well as spontaneous recovery potential.

C. Pre-stroke LFP do not contribute to motor recovery prediction

Our results showed that incorporating pre-stroke brain signals did not improve the prediction accuracy of motor recovery. This indicates that motor recovery is more closely linked to the aftermath of the stroke rather than preexisting brain dynamics. The lack of a significant effect of pre-stroke data suggests that any neuroprotective effects presented before the stroke may not be substantial enough to influence recovery outcomes.

Interestingly, even combining the information of both pre- and post-stroke signals did not enhance prediction accuracy. In fact, normalizing the post-stroke LFP features with their respective pre-stroke values resulted in reduced accuracy compared to using post-stroke features alone. This suggests that the post-stroke brain activity contains the most relevant information for predicting recovery, and that pre-stroke signals may introduce variability that confounds the predictive model. This highlights the predominant role of the immediate post-stroke period in determining recovery trajectories and underscores the importance of focusing on post-stroke brain dynamics for accurate predictions.

While pre-stroke electrophysiological signals may lack biomarkers for motor recovery, this does not rule out the potential presence of pre-stroke prognostic factors. For instance, motor enrichment and exercise before stroke have been linked to improved motor recovery outcomes.[50–52]

D. Translational potential from murine models to human studies

In this study, we demonstrated that LFPs in mice carry valuable information about future motor recovery following a stroke and can be effectively utilized to construct predictive models for recovery outcomes. The machine learning pipeline we developed and the features we extracted are readily transferable to other recording modalities, such as EEG.

Although direct application to human subjects is limited since LFP recordings from motor areas post-stroke are rarely feasible, we identified significant parallels with findings in human EEG research. Specifically, we observed that interhemispheric balance metrics, such as Granger causality from the contralesional to the ipsilesional hemisphere, serve as key predictors of recovery. Similar patterns have been reported in humans, where metrics like the pairwise-derived brain symmetry index and related measures have been shown to predict motor recovery.[24,29,30] Furthermore, consistent with human studies, ipsilesional relative band power does not seem to provide meaningful information about motor recovery,[24] with negative prognosis instead emphasized by increased contralesional activity of the low frequency bands.[53]

In addition, we explored novel features that have been less commonly investigated in human EEG research and found them to be predictive of recovery. For example, approximate entropy, a measure of signal complexity, emerged as a valuable predictor. These findings provide an opportunity to benchmark such features before testing them in human studies.

The translational relevance of this work is hence twofold. First, we developed a machine learning pipeline that can be easily adapted for human applications using different recording modalities. Second, our findings indicate that some of the extracted features may have broad applicability across species and recording types, positioning the mouse stroke model as a valuable platform for validating features with potential relevance to human recovery.

However, while our findings suggest that key LFP-derived features may translate to EEG, this requires empirical validation in human stroke patients. Future studies should determine whether these features retain prognostic value in human EEG data and assess their reliability as recovery biomarkers. This validation is a crucial and necessary step toward translating our machine learning framework to a noninvasive clinical setting.

E. Limitations and future perspectives

Here, we used a murine model of unilateral ischemic stroke. The MCAO model is widely used to study ischemic stroke, simulating the effects of reduced blood flow in the brain. While this model provides valuable insight into stroke mechanisms and potential therapies, it has the limitation of representing only a subset of human stroke cases. In fact, in our dataset, the ischemic lesion was relatively confined and essentially located in cortical layers without affecting subcortical regions.

Our study investigated the predictive value of features extracted from LFPs in the motor cortex of freely moving mice. LFPs were recorded from both ipsilateral and contralateral motor cortices using two electrodes. Expanding the number of electrode contacts could significantly enhance our findings. Specifically, using multi-contact electrodes that span the entire cortical thickness could help capture layer-specific differences, potentially improving the ability to distinguish between good and poor motor recovery following stroke. Additionally, multi-contact probes would potentially enable the inclusion of both single-unit and multi-unit activities, providing complementary information on neural dynamics compared to LFPs alone,[54] and thus, offering a more comprehensive understanding of post-stroke recovery.

Furthermore, we aimed to assess the predictive value of motor LFPs under the most natural and unconstrained conditions. To achieve this, we recorded extracellular potentials during freely moving behavior, allowing mice to engage in spontaneous forelimb activity with minimal experimental constraints, which better reflects self-determined behavior. However, since the severity of initial motor impairment remains one of the most reliable predictors of motor outcome,[55,56] task-evoked or stimuli-induced LFPs might reveal more pronounced differences in motor recovery following stroke, as they may more effectively indicate the integrity of the motor system in the aftermath of stroke. Combining resting state with evoked LFPs could, therefore, enhance the predictive value of such signals in the stroke aftermath.

A potential concern in the LFP recordings used in our work is variability introduced by minor differences in electrode placement across animals. However, we followed a standardized stereotaxic implantation procedure (see "Electrode implant"), and the spatial correlation of LFP signals extends over the millimeter scale,[57,58] hence mitigating the impact of small electrode placement mismatches. Furthermore, while acute tissue damage could influence invasive recordings, we implemented a post-operative recovery period, allowing neural activity to stabilize before data collection. Finally, the predictive features identified in our analysis were robust across animals, suggesting that our findings are not primarily driven by implantation variability.

Our analysis heavily relied on the selected features. We intentionally included features that are straightforward to use and have previously been shown to effectively capture stroke-related pathological dynamics in both mice and humans.[29,45,59] However, it is possible that we excluded certain features that could have improved the prediction accuracy in this study. To this end, a potential future direction for the approach proposed here involves the use of automatic feature extraction, such as employing deep learning techniques. These advanced methods could potentially identify and leverage additional relevant features that manual selection may overlook, thereby enhancing the predictive performance of the model here proposed.

While our methodological choices were carefully designed to mitigate the risk of overfitting, we acknowledge the inherent challenges posed by the limited dataset size ($n = 16$ animals). The use of LOAO cross-validation ensures that classification is assessed at the animal level, but it also reduces the amount of data available for training (by one animal) in each iteration. Although our results demonstrate that classification performance remains stable even after removing the least informative features (suggesting that the model is not overfitting due to excessive dimensionality) future studies with larger datasets would further validate our approach. Additionally, while we employed multiple cross-validation strategies and data augmentation via sliding-window to enhance generalization, further validation on independent datasets will be crucial to fully establish the robustness and broader applicability of our findings.

Another crucial factor influencing the predictive power is the definition of motor recovery itself. We utilized a motor recovery score system based on the combined result of multiple motor tests. This approach was chosen because it provides a more comprehensive definition of motor recovery following a stroke, capturing a broader spectrum of motor functions compared to single behavioral tests, which may focus on only one aspect of motor output and fail to reflect the complexity of real-world scenarios. However, it could be valuable to extend the proposed approach to predict motor recovery outcomes for

each individual motor test, rather than relying solely on an aggregate measure. Such an extension could potentially enhance the precision of predictions by providing insight into specific aspects of motor recovery, thereby offering a more detailed understanding of recovery dynamics and improving the overall accuracy of our predictive model.

A potential limitation of using a global motor score is that it aggregates multiple behavioral tests, which may obscure distinctions between different aspects of motor recovery. However, in our dataset, individual motor tests probed distinct motor skills and, as a result, classified animals differently as good or poor recoverers. This variability led to inconsistencies across tests, where some identified more animals as good or poor recoverers than others, making group assignments less stable and imbalanced. Given our sample size, addressing these inconsistencies through resampling techniques would not be ideal, as such methods could amplify the influence of individual animals rather than capture generalizable recovery patterns. By integrating multiple behavioral measures, the global motor score mitigated this issue, providing a more stable classification and a broader perspective on post-stroke motor recovery. Additionally, from a translational perspective, this approach better aligns with clinical assessments in human patients, which commonly rely on composite motor scales rather than isolated test outcomes. Future studies with larger sample sizes may refine this approach by investigating recovery within specific motor domains (e.g., fine vs gross motor recovery). A larger dataset would enable more stable groupings and allow for resampling techniques to correct for discrepancies across individual tests, potentially improving the resolution of motor recovery classification.

Notably, the model proposed here could be straightforwardly extended to address non-spontaneous motor recovery. Specifically, it could be expanded to test the possible existence of prognostic factors related to motor therapy. By incorporating these elements, our model could provide valuable predictions regarding the effectiveness of different types of therapeutic interventions, further enhancing its clinical applicability and utility in guiding personalized rehabilitation strategies.

Finally, our prediction model focused on features derived from electrophysiological signals. However, the prediction of motor recovery following stroke has primarily been associated with neuroimaging biomarkers,[60] particularly those assessing the integrity of corticospinal tract white matter.[61–63] Integrating electrophysiological data with neuroimaging features could potentially enhance prediction accuracy through a multimodal approach, combining both types of information for a more comprehensive assessment.

IV. METHODS

A. Animals

All experiments involving animals were authorized by the Italian Ministry of Health (authorization number 684/2020-PR) and carried out in accordance with the European Directive of 22 September 2010 (EU/63/2010). Mice were housed in clear plastic cages under a 12 h light/dark cycle and were given *ad libitum* access to water and food. A total of 16 C57BL6J mice were used (22–27 g, age 8–12 weeks).

B. Permanent middle cerebral artery occlusion (MCAO)

Animals were anesthetized with avertin (20 ml/kg, 2,2,2 tribromoethanol 1.25%; Sigma-Aldrich, USA), transferred onto a heat blanket in lateral position and immobilized with adhesive tape for medical use. A skin incision was performed between the ear and eye using fine surgical scissors. Then, the temporal muscle was exposed, detached from the bone, and cut vertically to reveal the underlying skull. The MCA was identified below the semitransparent bone and a circular craniotomy was drilled above it. Once it was exposed, the MCA was occluded with electrocoagulation forceps (two biological instruments, bipolar mode, 12 W). If the artery was bifurcated, the coagulation was performed both proximally and distally to the branching and, if necessary, also the single branches were coagulated. When there was no bifurcation, the MCA was coagulated at two sites of approximately 1 mm of distance. If spontaneous recanalization occurred, the electrocoagulation was performed again. Finally, the temporal muscle was relocated to its position, glued with cyanoacrylate and the wound was sutured. Intraoperative rectal temperature was kept at $37.0 \pm 0.5\,^{\circ}\mathrm{C}$ using a heating pad. After the surgical procedure, animals were treated with paracetamol (100 mg/kg) for 4 days post operation in drinking water.

C. Electrode implant

Animals were anesthetized with avertin (20 ml/kg, 2,2,2 tribromoethanol 1.25%; Sigma-Aldrich, USA), placed in a stereotaxic apparatus and the skull was exposed through a midline incision. Then, two burr holes were drilled in both hemispheres at the level of Bregma, 2 mm lateral. This way, the electrodes were implanted in the caudal forelimb area (CFA).[64,65] A surgical screw was tightened at the center of the occipital bone and was used both as ground reference and to give more stability to the recording implant. Bipolar electrodes were assembled prior to the surgery: a couple twitched insulated tungsten wires 700 μm spaced in length at the tip) were soldered at the channel and reference pins of the connector (Elitalia). With the aid of a micromanipulator, the electrode was stereotactically inserted in the center of the drilled holes at 700 μm of depth and the ground pin was connected to the occipital screw. A first layer of dental cement (Superbond, Sun Medical) was distributed to secure all the components to the skull surface. After the first layer of cement was dried, an aluminum post, used later to head-fix the animal, was cemented onto the skull. Finally, a second layer of cement (Paladur, Pala, Germany) was distributed to enclose all the electrical components. After the surgical procedure, animals were treated with paracetamol (100 mg/kg) for 4 days post-operation in drinking water. Animals were allowed to habituate to the implant for 1 week before undergoing any further procedures.

D. Global motor score

A battery of behavioral tests was conducted in order to assess whether a measurable motor deficit was induced by MCAO into a motor score. Motor performance was assessed in baseline condition (i.e., pre-stroke) and 30 days post lesion [Fig. 1(a)] using Gridwalk, Skilled Reaching Test and retraction task in the robotic M-platform.

Gridwalk Test: animals were allowed to walk freely for 5 min on an elevated grid (32×20 cm, with 11×11 mm-large openings) and the task was video-recorded. In order to quantify forelimb-use asymmetry displayed by the animal, an Asymmetry Index was computed, as previously described.[66]

Skilled Reaching Test: The percentage of correct movements and the kinematic analysis of the whole reaching movements were performed, according to Lai *et al.*[66] Briefly, animals (food deprived for 15 h) were placed in a testing chamber with plastic walls and trained to

perform a skilled reaching task with their preferred paw, which had to pass through a small frontal rectangular aperture (0.5 × 1.3 cm) to grasp and retrieve food pellets. The task was recorded by video camera, placed on the side of the testing chamber, thus allowing for a sagittal view of the animal. The number of correct (i.e., a reach and grasp movements ending with pellet eating) and incorrect movements (i.e., when the mouse passed by the frontal window and reached the pellet but either missed it or dropped the pellet after grasping it) were manually assessed. Then, the percentage of incorrect grasping was calculated on the total attempts, defined as every time that the paw crossed the frontal window.

Retraction Test in the Robotic M-platform: We measured forelimb performances in a robotic device, the M-Platform,[67,68] where the animal had to pull back a handle in order to get a liquid reward thereby allowing the assessment of kinetic and kinematic parameters. From these kinematic and kinetic signals, a series of parameters were automatically computed to describe detailed motor performance on the platform including the t-target (i.e., the time spent by the animal to accomplish a single retraction task) and the number of attempts to move the handle, with "attempts" defining the force exerted by the mouse not sufficient to overcome the static friction. The computation and statistical analysis of these parameters were performed using custom-made algorithms developed in Matlab (Mathwork, USA).[68]

To calculate the global motor score (GMS), the readouts of each behavioral test were converted in a scale from 0 to 10. Score 0 was assigned to animals showing a bad performance, while score 10 to represents maximal performance. The converted readouts were summed up to obtain the global motor score. Mice with a 30-day GMS lower than their baseline GMS – 1SD (measured under baseline conditions) were classified as poor recoverers (n = 8), while all others as good recoverers (n = 8).

E. LFPs acquisition

Spontaneous neuronal activity was recorded in N = 16 animals for 30 min during freely moving behavior before surgery and then at 2- and 30-days post lesion [Fig. 1(a)]. The animals were placed in a recording chamber, where, after a 1-h habituation, local field potential (LFP) recording sessions were performed. Signals were acquired by a miniature headstage (NPI, Germany) connected to an amplifier (EXT-02F, NPI). Signals were amplified (10.000-fold), filtered (high pass 0.3 Hz; low pass, 100 Hz), digitized with a sampling rate fs = 200 Hz (National Instruments Card), and conveyed to a computer for a storage and analysis.

F. LFPs pre-processing

To detect artifacts caused by sudden or jerky movements, we used a peak detection algorithm (findpeaks in Matlab) on the extracellular signal envelope, as in Refs. 45 and 69. The envelopes were calculated using Matlab's envelope function, which determines the root mean square of the signal with a 1-s sliding window. Local maxima were identified using a peak detection algorithm, with a minimum amplitude threshold set at three times the signal's standard deviation.

For each animal, we selected the longest segments of recordings free from artifacts for both ipsilesional and contralesional LFPs, with a maximum duration of 600 s. This criterion was met in 13 out of 16 mice. In the remaining 3 mice, the longest artifact-free recording segment was 500 s.

G. Windowed LFPs feature extraction

To increase the dataset size, we extracted features from the LFPs using sliding windows [Fig. 1(b)]. We selected two window lengths (30 and 60 s) and employed three overlapping strategies: (i) 50% overlap between consecutive windows, (ii) no overlap between consecutive windows, and (iii) a random overlap fraction between consecutive windows.

For each temporal window we extracted a set of 28 features grouped into three categories: 8 based on spectral power (for each hemisphere); 4 based on signal complexity (for each hemisphere); and 4 accessing inter-hemispheric interaction.

H. Power features

We extracted from each LFP window, both in ipsilesional and contralesional hemispheres, a set of 8 features based on the signal spectral patterns [Fig. 1(b)].

We extracted the power of the LFP windows in five different bands: delta (or δ) [0.5–4] Hz; theta (or θ) [4–8] Hz; alpha (or α) [8–12] Hz; beta (or β) [12–30] Hz; gamma (or γ) [30–50] Hz. The power was obtained as the integral of the power spectral density (PSD) (obtained with the Welch method in a 1 s long 50% overlapping sliding windows) in each band of interest, normalized by the signal total power.

Previous reports have shown that phase-amplitude cross-frequency coupling (cfc) can serve as a cortical biomarker for probing motor recovery following ischemic lesions.[70,71] Consequently, we computed the phase-amplitude cfc between the phase within the delta and theta bands and the gamma power, referred to in Fig. 1(b) as "cfc delta/gamma" and "cfc theta/gamma."

Phase and power dynamics were estimated using the Hilbert transform on bandpass filtered LFP windows within the bands of interest. The phase-amplitude cfc was defined as the maximum-minimum difference in gamma power average amplitudes across discrete delta/theta phase bins.

Additionally, the slope of the aperiodic component of the power spectral density (PSD) has been shown to track fine motor recovery following ischemic lesions in mice.[72] We, therefore, estimated the aperiodic slope of the PSD, termed "aper. PSD slope," using the FOOOF toolbox.[73]

I. Complexity features

Estimates of brain complexity have been shown in various studies to track electrophysiological hallmarks of ischemic lesions and motor recovery.[44,74–79]

From each LFP window, both in the ipsilesional and contralesional hemispheres, we extracted a set of four features based on LFP complexity [Fig. 1(b)].

Approximate entropy and range sample entropy are non-parametric methods to estimate the randomness of time series. For a detailed mathematical foundation of these algorithms, interested readers are referred to Omidvarnia *et al.*[80] In our context, we estimated these metrics using (i) the built-in Matlab function "approximateEntropy"

and (ii) the freely accessible implementation of the range sample entropy algorithm available from Palmer *et al.*[81]

The Hurst exponent is used to determine the degree of long-term memory and predictability in a time series, identifying whether the series exhibits persistent, anti-persistent, or random behavior over time. We utilized the freely available implementation presented here: https://it.mathworks.com/matlabcentral/fileexchange/9842-hurst-exponent.

Finally, we estimated the Kolmogorov complexity, which quantifies the randomness and compressibility of a data string. For this purpose, we used the implementation available at https://it.mathworks.com/matlabcentral/fileexchange/6886-kolmogorov-complexity.

J. Interhemispheric interaction features

Previous studies have demonstrated that the impairment, strength, and characteristics of interhemispheric interactions are crucial in capturing post-stroke motor function and recovery.[45,82–84] Accordingly, we assessed the interaction between ipsilesional and contralesional hemispheres using four metrics: Pearson's linear correlation coefficient, mutual information, and Granger causality test statistics of contralateral on ipsilesional LFPs, and vice versa.

Pearson's linear correlation coefficient was estimated using the "corr" MATLAB function. Mutual information was applied to LFP data, as in previous works from our group,[85,86] using the freely available toolbox presented in Magri *et al.*[87] Briefly, contralateral and ipsilateral signals were discretized using 10 equi-populated bins, and mutual information was estimated. Bias was accounted for by using the Panzeri–Treves correction.[88]

Finally, Granger causality tests were performed on the selected best vector autoregressive (VAR) model to determine the predictive causality between ipsilateral and contralateral LFPs. Specifically, we first fitted a VAR model with temporal lags ranging from 1 to 10 samples. The best fit was identified by minimizing the Akaike Information Criterion across lags. For both hemispheres' LFPs, we conducted a Granger causality test using the MATLAB function gctest to assess whether one hemisphere's LFP in the fitted VAR model is a 1-step Granger-cause of the other hemisphere's. The strength of the causal relationship between the two time series was defined as the test statistic output from the Granger causality test, which in this case was derived from conducting a chi-squared (χ^2) test.

After concatenating all the features, the input feature space class separation was visualized in two dimensions using t-SNE.[49]

K. Classification experiments

To automatically predict which animals will recover based on LFP features, we conducted three classification experiments. In each experiment, we aimed to predict the recovery or non-recovery status of each animal at 30 days post-stroke, according to the motor recovery score described in Sec. IV D. First, we used LFP features computed at 2 days post-stroke as inputs for training the algorithms (experiment 1). This approach is based on the hypothesis that functional motor recovery in the chronic phase can be predicted from measures obtained in the acute phase.[23,28,59] Second, we normalized the features computed at 2 days post-stroke by their corresponding baseline (pre-stroke) values (experiment 2). In the third experiment, we used the baseline pre-stroke features to predict recovery at 30 days post-stroke (experiment 3). The goal of the last two experiments was to explore whether pre-stroke information could be used to predict which animals would recover.

L. Training-validation strategies and classification performance

Classification experiments were conducted using a nested leave-one-animal-out cross-validation (LOAO-CV) approach. This means that in each iteration, data from one animal (i.e., LFP features extracted from multiple time windows) was left out as an independent test set (outer LOAO-CV). For the remaining animals, an inner LOAO-CV was performed: one animal's data at a time was excluded, and the classification algorithms were trained using all possible combinations of hyperparameters, window lengths, and overlaps to identify the best combination for maximizing classification performance. The optimal set of training parameters was then used to train the classification model on the inner LOAO-CV dataset. This trained model was subsequently tested on the animal left out in the outer LOAO-CV. This process ensured that for each animal we conducted a comprehensive grid search of hyperparameters to develop a personalized classification model. We chose this nested LOAO-CV strategy over a fixed split or stratified k-fold cross-validation due to the limited number of animals. This approach minimizes performance fluctuations that can arise from the variability in the training and test sets, ensuring more reliable and robust results.

Classification models were compared in the inner LOAO-CV using the classification accuracy to select the best combination of training parameters. Accuracy, precision, recall, and F1-score were used to evaluate the classification performance on the test animal in the outer LOAO-CV.

Since the classification algorithms were trained on LFP features extracted from multiple windows, they provided a prediction for each temporal window of the test animal. Consequently, to obtain a prediction at the animal level, we assigned a label (recoverer or non-recoverer) to each animal based on the majority vote from its respective windows.

The training-validation strategy for the baseline LFPs (experiments 2 and 3) required additional adjustments, as the classification accuracy with LOAO-CV dropped significantly below chance (see Sec. II).

To address this issue, we performed a leave-two-animals-out cross-validation (L2AO-CV) procedure. Specifically, for every animal iteratively assigned to the test set, we also iteratively assigned one animal from the opposite class to the same test set, in an exhaustive manner. This ensured that, at each iteration, the training/validation set was balanced.

The use of the L2AO-CV approach was necessary due to the interaction between the LOAO testing procedure and the label assignment based on majority voting across windows. Specifically, as the baseline LFPs dataset lacked informative content relevant to the classification task, the combination of the LOAO procedure with majority voting for label assignment led to biased prediction labels. These labels were skewed toward the class opposite to that of the test subject (which with the LOAO became the majority class), ultimately resulting in classification accuracy falling below chance level.

We further tested the prediction of motor recovery using conventional fixed-split cross-validation (fixed-CV, Fig. S1). We split the dataset into two sets: a test set composed of six animals (three randomly

selected good recoverers and three randomly selected bad recoverers), while the remaining animals comprised the training/validation set. Hyperparameters were optimized based on the LOAO animal-wise accuracy on the training/validation set.

M. Classification algorithms

The classification algorithms used in this study were support vector machine (SVM), random forest (RF), feed-forward fully connected neural network (NN), and k-nearest neighbor (kNN). To optimize their performance, we conducted a grid search over multiple hyperparameters within the inner loop of our nested leave-one-animal-out cross-validation (LOAO-CV). This approach ensured that the best set of hyperparameters was selected in an unbiased manner while accounting for inter-animal variability (see "Training-validation strategies and classification performance"). For SVM, we explored different kernel types (radial basis function—RBF, linear, quadratic, cubic functions) and systematically varied the C and gamma parameters using a logarithmic scale from 0.001 to 1000 in seven steps. For RF, we tested different numbers of trees: 10, 20, 50, 100, 200, 500, and 1000. For NN, we considered all possible combinations of architectures with 1, 2, and 3 layers, each containing 5, 10, 15, 20, or 25 nodes per layer. For kNN, we tested values of $k = 3, 5, 7$, and 9.

Additionally, all hyperparameter searches were conducted while systematically varying the windowing strategy (no overlap, 50% overlap, random overlap) and window duration (30 or 60 s).

Hyperparameters were selected based on the highest animal-wise classification accuracy in the inner LOAO-CV (i.e., the number of correctly classified animals). In cases where multiple hyperparameter sets led to the same animal-wise classification accuracy, the best hyperparameters were chosen as those maximizing the classification accuracy at the window level. The optimal set of hyperparameters was then used to train the final model, which was evaluated on the left-out animal in the outer LOAO-CV. This nested approach helped mitigate the risk of overfitting while ensuring robust generalization given the limited sample size.

N. Feature selection

In addition to conducting classification experiments using the full set of features (i.e., all 28 LFP features described in Secs. IV H—IV J), we also evaluated the classification performance using only a subset of these features.

We used the maximum relevance minimum redundancy (mRMR) algorithm[89] to rank the feature importance. Briefly, mRMR is a minimal-optimal feature selection algorithm. This means it iteratively ranks the relevance of features for a given machine learning task while minimizing redundancy with previously identified features.

In our case, we applied the mRMR algorithm together with the LOAO-CV testing procedure. Each LOAO-CV iteration (where each animal iteratively served as the test animal) resulted in a different and unique feature ranking order. To aggregate the feature importance results, we reported the average ranking of the features in the manuscript [Fig. 3(a)].

In this work, we used the Python implementation of the mRMR algorithm freely available here: https://github.com/smazzanti/mrmr.

O. Statistical analysis

All statistical analyses were performed in MATLAB (MathWorks, MA). Results and details of the statistical comparisons performed in the study including sample sizes were reported in the results sections and figure legends.

Two-sided Wilcoxon rank-sum test was used to compare the features distributions between good and poor motor recoverer and between pre- and post-stroke condition. A mixed-design ANOVA was employed to evaluate the effects of two factors: stroke (baseline vs 2 days post-stroke, i.e., the within subject factor) and group (good vs poor recoverers, i.e., the between subject factor), along with their interaction.

SUPPLEMENTARY MATERIAL

See the supplementary material for figures reporting (1) the accuracy of motor recovery prediction using fixed-split CV and (2) animal-wise average of the significantly different features between good and poor motor recoverers.

ACKNOWLEDGMENTS

This work was supported by the PERSONA Project, which is funded by the Government of Tuscany (Italy)—Bando Salute 2018. This work was also supported by #NEXTGENERATIONEU (NGEU) and funded by the Ministry of University and Research (MUR), National Recovery and Resilience Plan (NRRP), project MNESYS (PE0000006)—A Multiscale integrated approach to the study of the nervous system in health and disease (DN. 1553 11.10.2022) and project THE (IECS00000017) - Tuscany Health Ecosystem (DN. 1553 11.10.2022). The work was also funded by "Fondo di beneficenza ed opere di carattere sociale e culturale" granted by "Banca Intesa San Paolo" in the cotext of the project ONDA ("Origini neurali delle disfunzioni dell'arto superiore su modello murino di parkinson"), CUP B53C22007580007, and BIOMETRICS (Nuovi biomarcatori elettrofisiologici predittivi del recupero post-ictus per una terapia personalizzabile), project (No. B/2023/0232).

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Ethics Approval

Ethics approval for experiments reported in the submitted manuscript on animal or human subjects was granted. Ethics approval for experiments reported in the submitted manuscript on animals was granted. All experiments involving animals were authorized by the Italian Ministry of Health (authorization number 684/2020-PR) and carried out in accordance with the European Directive of 22 September 2010 (EU/63/2010).

Author Contributions

Michael Lassi and Verediana Massa contributed equally to this work.

Nicolò Meneghetti: Investigation (equal); Methodology (lead); Software (equal); Visualization (lead); Writing – original draft (lead);

Writing – review & editing (equal). **Michael Lassi:** Methodology (supporting); Software (supporting); Writing – original draft (equal); Writing – review & editing (equal). **Verediana Massa:** Investigation (equal); Methodology (equal); Writing – original draft (equal); Writing – review & editing (supporting). **Silvestro Micera:** Conceptualization (equal); Funding acquisition (equal); Writing – review & editing (equal). **Alberto Mazzoni:** Conceptualization (equal); Funding acquisition (equal); Writing – review & editing (equal). **Claudia Alia:** Conceptualization (equal); Funding acquisition (equal); Methodology (equal); Supervision (equal); Writing – original draft (equal); Writing – review & editing (equal). **Andrea Bandini:** Conceptualization (equal); Methodology (equal); Software (equal); Supervision (equal); Writing – original draft (equal); Writing – review & editing (equal).

DATA AVAILABILITY

The data that support the findings of this study are openly available in GitHub, Ref. 90.

REFERENCES

[1]V. Saini, L. Guada, and D. R. Yavagal, "Global epidemiology of stroke and access to acute ischemic stroke interventions," Neurology **97**, S6–S16 (2021).
[2]S. Anwer, A. Waris, S. O. Gilani, J. Iqbal, N. Shaikh, A. N. Pujari, and I. K. Niazi, "Rehabilitation of upper limb motor impairment in stroke: A narrative review on the prevalence, risk factors, and economic statistics of stroke and state of the art therapies," Healthcare **10**, 190 (2022).
[3]V. L. Feigin, B. A. Stark, C. O. Johnson, G. A. Roth, C. Bisignano, G. G. Abady, M. Abbasifard, M. Abbasi-Kangevari, F. Abd-Allah, V. Abedi *et al.*, "Global, regional, and national burden of stroke and its risk factors, 1990–2019: A systematic analysis for the Global Burden of Disease Study 2019," Lancet Neurol. **20**, 795–820 (2021).
[4]S. Li, "Spasticity, motor recovery, and neural plasticity after stroke," Front. Neurol. **8**, 120 (2017).
[5]S. T. Carmichael, "Plasticity of cortical projections after stroke," Neuroscientist **9**, 64–75 (2003).
[6]B. Hordacre, D. Austin, K. E. Brown, L. Graetz, I. Pareés, S. De Trane, A.-M. Vallence, S. Koblar, T. Kleinig, M. N. McDonnell *et al.*, "Evidence for a window of enhanced plasticity in the human motor cortex following ischemic stroke," Neurorehabil. Neural Repair **35**, 307–320 (2021).
[7]C. Alia, C. Spalletti, S. Lai, A. Panarese, G. Lamola, F. Bertolucci, F. Vallone, A. Di Garbo, C. Chisari, S. Micera *et al.*, "Neuroplastic changes following brain ischemia and their contribution to stroke recovery: Novel approaches in neurorehabilitation," Front. Cell. Neurosci. **11**, 76 (2017).
[8]M. Caleo, "Rehabilitation and plasticity following stroke: Insights from rodent models," Neuroscience **311**, 180–194 (2015).
[9]R. J. de Haan, M. Limburg, J. H. P. Van der Meulen, H. M. Jacobs, and N. K. Aaronson, "Quality of life after stroke," Stroke **26**, 402–408 (1995).
[10]C. Kugler, T. Altenhöner, P. Lochner, A. Ferbert, H. for the, D. Stroke, S. Bank, and A. S. H. Group, "Does age influence early recovery from ischemic stroke?," J. Neurol. **250**, 676–681 (2003).
[11]M. Kotila, O. Waltimo, M. L. Niemi, R. Laaksonen, and M. Lempinen, "The profile of recovery from stroke and factors influencing outcome," Stroke **15**, 1039–1044 (1984).
[12]L. Jongbloed, "Prediction of function after stroke: A critical review," Stroke **17**, 765–776 (1986).
[13]J. W. Krakauer and R. S. Marshall, "The proportional recovery rule for stroke revisited," Ann. Neurol. **78**, 845–847 (2015).
[14]R. Kundert, J. Goldsmith, J. M. Veerbeek, J. W. Krakauer, and A. R. Luft, "What the proportional recovery rule is (and is not): Methodological and statistical considerations," Neurorehabil. Neural Repair **33**, 876–887 (2019).
[15]C. Winters, E. E. H. van Wegen, A. Daffertshofer, and G. Kwakkel, "Generalizability of the proportional recovery model for the upper extremity after an ischemic stroke," Neurorehabil. Neural Repair **29**, 614–622 (2015).
[16]T. M. H. Hope, K. Friston, C. J. Price, A. P. Leff, P. Rotshtein, and H. Bowman, "Recovery after stroke: Not so proportional after all?," Brain **142**, 15–22 (2019).
[17]C. M. Stinear, "Prediction of motor recovery after stroke: Advances in biomarkers," Lancet Neurol. **16**, 826–836 (2017).
[18]C. M. Stinear, W. D. Byblow, S. J. Ackerley, M. Smith, V. M. Borges, and P. A. Barber, "PREP2: A biomarker-based algorithm for predicting upper limb function after stroke," Ann. Clin. Transl. Neurol. **4**, 811–820 (2017).
[19]M. Lotze, W. Beutling, M. Loibl, M. Domin, T. Platz, U. Schminke, and W. D. Byblow, "Contralesional motor cortex activation depends on ipsilesional corticospinal tract integrity in well-recovered subcortical stroke patients," Neurorehabil. Neural Repair **26**, 594–603 (2012).
[20]M. A. Perez and L. G. Cohen, "The corticospinal system and transcranial magnetic stimulation in stroke," Top. Stroke Rehabil. **16**, 254–269 (2009).
[21]V. Longo, S. A. Barbati, A. Re, F. Paciello, M. Bolla, M. Rinaudo, F. Miraglia, F. Alù, M. G. Di Donna, F. Vecchio *et al.*, "Transcranial direct current stimulation enhances neuroplasticity and accelerates motor recovery in a stroke mouse model," Stroke **53**, 1746–1758 (2022).
[22]D. S. Ramanathan, L. Guo, T. Gulati, G. Davidson, A. K. Hishinuma, S.-J. Won, R. T. Knight, E. F. Chang, R. A. Swanson, and K. Ganguly, "Low-frequency cortical activity is a neuromodulatory target that tracks recovery after stroke," Nat. Med. **24**, 1257–1267 (2018).
[23]A. A. Vatinno, A. Simpson, V. Ramakrishnan, H. S. Bonilha, L. Bonilha, and N. J. Seo, "The prognostic utility of EEG in stroke recovery: A systematic review and meta-analysis," Neurorehabil. Neural Repair **36**, 255–268 (2022).
[24]M. Saes, C. G. M. Meskers, A. Daffertshofer, J. C. de Munck, G. Kwakkel, E. E. H. van Wegen, and 4D-EEG consortium, "How does upper extremity Fugl-Meyer motor score relate to resting-state EEG in chronic stroke? A power spectral density analysis," Clin. Neurophysiol. **130**, 856–862 (2019).
[25]G. Milani, A. Antonioni, A. Baroni, P. Malerba, and S. Straudi, "Relation between EEG measures and upper limb motor recovery in stroke patients: A scoping review," Brain Topogr. **35**, 651 (2022).
[26]S. Finnigan, A. Wong, and S. Read, "Defining abnormal slow EEG activity in acute ischaemic stroke: Delta/alpha ratio as an optimal QEEG index," Clin. Neurophysiol. **127**, 1452–1459 (2016).
[27]J. Leon-Carrion, J. F. Martin-Rodriguez, J. Damas-Lopez, J. M. Barroso y Martin, and M. R. Dominguez-Morales, "Delta-alpha ratio correlates with level of recovery after neurorehabilitation in patients with acquired brain injury," Clin. Neurophysiol. **120**, 1039–1045 (2009).
[28]M. Saes, S. B. Zandvliet, A. S. Andringa, A. Daffertshofer, J. W. R. Twisk, C. G. M. Meskers, E. E. H. van Wegen, and G. Kwakkel, "Is resting-state EEG longitudinally associated with recovery of clinical neurological impairments early poststroke? A prospective cohort study," Neurorehabil. Neural Repair **34**, 389–402 (2020).
[29]M. Saes, C. G. M. Meskers, A. Daffertshofer, E. E. H. van Wegen, and G. Kwakkel, "Are early measured resting-state EEG parameters predictive for upper limb motor impairment six months poststroke?," Clin. Neurophysiol. **132**, 56–62 (2021).
[30]A. Agius Anastasi, O. Falzon, K. Camilleri, M. Vella, and R. Muscat, "Brain symmetry index in healthy and stroke patients for assessment and prognosis," Stroke Res. Treat. **2017**, 276136.
[31]M. J. A. M. van Putten and D. L. J. Tavy, "Continuous quantitative EEG monitoring in hemispheric stroke patients using the brain symmetry index," Stroke **35**, 2489–2492 (2004).
[32]M. Lassi, S. Dalise, A. Bandini, V. Spina, V. Azzollini, M. Vissani, S. Micera, A. Mazzoni, and C. Chisari, "Neurophysiological underpinnings of an intensive protocol for upper limb motor recovery in subacute and chronic stroke patients," Eur. J. Phys. Rehabil. Med. **60**, 13 (2024).
[33]C. Fanciullacci, A. Panarese, V. Spina, M. Lassi, A. Mazzoni, F. Artoni, S. Micera, and C. Chisari, "Connectivity measures differentiate cortical and subcortical sub-acute ischemic stroke patients," Front. Hum. Neurosci. **15**, 669915 (2021).
[34]S. Eldeeb, M. Akcakaya, M. Sybeldon, S. Foldes, E. Santarnecchi, A. Pascual-Leone, and A. Sethi, "EEG-based functional connectivity to analyze motor recovery after stroke: A pilot study," Biomed. Signal Process. Control **49**, 419–426 (2019).
[35]A. Desowska and D. L. Turner, "Dynamics of brain connectivity after stroke," Rev. Neurosci. **30**, 605–623 (2019).

[36]F. Vecchio, C. Pappalettera, F. Miraglia, G. Deinite, R. Manenti, E. Judica, P. Caliandro, and P. M. Rossini, "Prognostic role of hemispherical functional connectivity in stroke: A study via graph theory versus coherence of electroencephalography rhythms," Stroke **54**, 499–508 (2023).
[37]M. Rubega, E. Formaggio, F. Molteni, E. Guanziroli, R. Di Marco, C. Baracchini, M. Ermani, N. S. Ward, S. Masiero, and A. Del Felice, "EEG fractal analysis reflects brain impairment after stroke," Entropy **23**, 592 (2021).
[38]G. Buzsáki, C. A. Anastassiou, and C. Koch, "The origin of extracellular fields and currents — EEG, ECoG, LFP and spikes," Nat. Rev. Neurosci. **13**, 407–420 (2012).
[39]N. Meneghetti, E. Vannini, and A. Mazzoni, "Rodents' visual gamma as a biomarker of pathological neural conditions," J. Physiol. **602**, 1017–1048 (2024).
[40]Y. Hara, "Brain plasticity and rehabilitation in stroke patients," J. Nippon Med. Sch. **82**, 4–13 (2015).
[41]N. Takeuchi and S.-I. Izumi, "Maladaptive plasticity for motor recovery after stroke: Mechanisms and approaches," Neural Plast. **2012**, 359728.
[42]J. L. Chen, M. L. Andermann, T. Keck, N.-L. Xu, and Y. Ziv, "Imaging neuronal populations in behaving rodents: Paradigms for studying neural circuits underlying behavior in the mammalian cortex," J. Neurosci. **33**, 17631–17640 (2013).
[43]M. Hazime, M. Alasoadura, R. Lamtahri, P. Quilichini, J. Leprince, D. Vaudry, and J. Chuquet, "Prolonged deficit of low gamma oscillations in the peri-infarct cortex of mice after stroke," Exp. Neurol. **341**, 113696 (2021).
[44]F. Miraglia, C. Pappalettera, S. A. Barbati, M. V. Podda, C. Grassi, P. M. Rossini, and F. Vecchio, "Brain complexity in stroke recovery after bihemispheric transcranial direct current stimulation in mice," Brain Commun. **6**, fcae137 (2024).
[45]F. Vallone, S. Lai, C. Spalletti, A. Panarese, C. Alia, S. Micera, M. Caleo, and A. D. Garbo, "Post-stroke longitudinal alterations of inter-hemispheric correlation and hemispheric dominance in mouse pre-motor cortex," PLoS One **11**, e0146858 (2016).
[46]A. M. Chiarelli, P. Croce, G. Assenza, A. Merla, G. Granata, N. M. Giannantoni, V. Pizzella, F. Tecchio, and F. Zappasodi, "Electroencephalography-derived prognosis of functional recovery in acute stroke through machine learning approaches," Int. J. Neural Syst. **30**, 2050067 (2020).
[47]M.-P. Hosseini, A. Hosseini, and K. Ahi, "A review on machine learning for EEG signal processing in bioengineering," IEEE Rev. Biomed. Eng. **14**, 204–218 (2021).
[48]M. Lassi, A. Bandini, V. Spina, V. Azzollini, S. Dalise, A. Mazzoni, C. Chisari, and S. Micera, "Classification of upper limb impairment in acute stroke patients using resting-state EEG markers and machine learning," in *2023 11th International IEEE/EMBS Conference on Neural Engineering (NER)* (IEEE, 2023), pp. 1–4.
[49]L. Van der Maaten and G. Hinton, "Visualizing data using t-SNE," J. Mach. Learn. Res. **9**, 2579–2605 (2008).
[50]M. R. Islam, M. F. Young, and C. D. Wrann, "Neuroprotective potential of exercise preconditioning in stroke," Cond. Med. **1**, 27–34 (2017).
[51]S. Hafez, Z. Eid, S. Alabasi, Y. Darwiche, S. Channaoui, and D. C. Hess, "Mechanisms of preconditioning exercise-induced neurovascular protection in stroke," J. Stroke **23**, 312–326 (2021).
[52]J. A. Kleim, T. A. Jones, and T. Schallert, "Motor enrichment and the induction of plasticity before or after brain injury," Neurochem. Res. **28**, 1757–1769 (2003).
[53]G. Assenza, F. Zappasodi, P. Pasqualetti, F. Vernieri, and F. Tecchio, "A contralesional EEG power increase mediated by interhemispheric disconnection provides negative prognosis in acute stroke," Restor. Neurol. Neurosci. **31**, 177–188 (2013).
[54]N. K. Logothetis, "The neural basis of the blood-oxygen-level-dependent functional magnetic resonance imaging signal," Philos. Trans. R. Soc. London, Ser. B **357**, 1003–1037 (2002).
[55]S. Prabhakaran, E. Zarahn, C. Riley, A. Speizer, J. Y. Chong, R. M. Lazar, R. S. Marshall, and J. W. Krakauer, "Inter-individual variability in the capacity for motor recovery after ischemic stroke," Neurorehabil. Neural Repair **22**, 64–71 (2008).
[56]F. Coupar, A. Pollock, P. Rowe, C. Weir, and P. Langhorne, "Predictors of upper limb recovery after stroke: A systematic review and meta-analysis," Clin. Rehabil. **26**, 291–313 (2012).
[57]A. Destexhe, D. Contreras, and M. Steriade, "Spatiotemporal analysis of local field potentials and unit discharges in cat cerebral cortex during natural wake and sleep states," J. Neurosci. **19**, 4595–4608 (1999).
[58]I. Nauhaus, L. Busse, M. Carandini, and D. L. Ringach, "Stimulus contrast modulates functional connectivity in visual cortex," Nat. Neurosci. **12**, 70–76 (2009).
[59]S. Finnigan and M. J. A. M. van Putten, "EEG in ischaemic stroke: Quantitative EEG can uniquely inform (sub-)acute prognoses and clinical management," Clin. Neurophysiol. **124**, 10–19 (2013).
[60]F. Knab, S. P. Koch, S. Major, T. D. Farr, S. Mueller, P. Euskirchen, M. Eggers, M. T. C. Kuffner, J. Walter, D. Berchtold *et al.*, "Prediction of stroke outcome in mice based on noninvasive MRI and behavioral testing," Stroke **54**, 2895–2905 (2023).
[61]B. N. Groisser, W. A. Copen, A. B. Singhal, K. K. Hirai, and J. D. Schaechter, "Corticospinal tract diffusion abnormalities early after stroke predict motor outcome," Neurorehabil. Neural Repair **28**, 751–760 (2014).
[62]J. M. Cassidy, G. Tran, E. B. Quinlan, and S. C. Cramer, "Neuroimaging identifies patients most likely to respond to a restorative stroke therapy," Stroke **49**, 433–438 (2018).
[63]J. Puig, G. Blasco, G. Schlaug, C. M. Stinear, P. Daunis-i-Estadella, C. Biarnes, J. Figueras, J. Serena, M. Hernández-Pérez, A. Alberich-Bayarri *et al.*, "Diffusion tensor imaging as a prognostic biomarker for motor recovery and rehabilitation after stroke," Neuroradiology **59**, 343–351 (2017).
[64]K. A. Tennant, D. L. Adkins, N. A. Donlan, A. L. Asay, N. Thomas, J. A. Kleim, and T. A. Jones, "The organization of the forelimb representation of the C57BL/6 mouse motor cortex as defined by intracortical microstimulation and cytoarchitecture," Cereb. Cortex **21**, 865–876 (2011).
[65]C. Alia, C. Spalletti, S. Lai, A. Panarese, S. Micera, and M. Caleo, "Reducing GABAA-mediated inhibition improves forelimb motor function after focal cortical stroke in mice," Sci. Rep. **6**, 37823 (2016).
[66]S. Lai, A. Panarese, C. Spalletti, C. Alia, A. Ghionzoli, M. Caleo, and S. Micera, "Quantitative kinematic characterization of reaching impairments in mice after a stroke," Neurorehabil. Neural Repair **29**, 382–392 (2015).
[67]C. Spalletti, C. Alia, S. Lai, A. Panarese, S. Conti, S. Micera, and M. Caleo, "Combining robotic training and inactivation of the healthy hemisphere restores pre-stroke motor patterns in mice," eLife **6**, e28662 (2017).
[68]C. Spalletti, S. Lai, M. Mainardi, A. Panarese, A. Ghionzoli, C. Alia, L. Gianfranceschi, C. Chisari, S. Micera, and M. Caleo, "A robotic system for quantitative assessment and poststroke training of forelimb retraction in mice," Neurorehabil. Neural Repair **28**, 188–196 (2014).
[69]J. H. Choi, K. P. Koch, W. Poppendieck, M. Lee, and H.-S. Shin, "High resolution electroencephalography in freely moving mice," J. Neurophysiol. **104**, 1825–1834 (2010).
[70]J. I. Mark, J. Riddle, R. Gangwani, B. Huang, F. Fröhlich, and J. M. Cassidy, "Cross-frequency coupling as a biomarker for early stroke recovery," Neurorehabil. Neural Repair **38**, 506–517 (2024).
[71]N. Rustamov, J. Humphries, A. Carter, and E. C. Leuthardt, "Theta–gamma coupling as a cortical biomarker of brain–computer interface-mediated motor recovery in chronic stroke," Brain Commun. **4**, fcac136 (2022).
[72]J. Biskamp, S. Isla Cainzos, F. L. Higgen, C. Gerloff, and T. Magnus, "Normalization of aperiodic electrocorticography components indicates fine motor recovery after sensory cortical stroke in mice," Stroke **53**, 2945–2953 (2022).
[73]M. Gerster, G. Waterstraat, V. Litvak, K. Lehnertz, A. Schnitzler, E. Florin, G. Curio, and V. Nikulin, "Separating neural oscillations from aperiodic 1/f activity: Challenges and recommendations," Neuroinformatics **20**, 991–1012 (2022).
[74]L. Liang, R. Hu, X. Luo, B. Feng, W. Long, and R. Song, "Reduced complexity in stroke with motor deficits: A resting-state fMRI study," Neuroscience **434**, 35–43 (2020).
[75]N. K. Al-Qazzaz, A. A. Aldoori, S. H. B. M. Ali, S. A. Ahmad, A. K. Mohammed, and M. I. Mohyee, "EEG signal complexity measurements to enhance BCI-based stroke patients' rehabilitation," Sensors **23**, 3889 (2023).
[76]N. K. Al-Qazzaz, S. H. B. M. Ali, S. A. Ahmad, M. S. Islam, and J. Escudero, "Discrimination of stroke-related mild cognitive impairment and vascular dementia using EEG signal analysis," Med. Biol. Eng. Comput. **56**, 137–157 (2018).
[77]S. Liu, J. Guo, J. Meng, Z. Wang, Y. Yao, J. Yang, H. Qi, and D. Ming, "Abnormal EEG complexity and functional connectivity of brain in patients

with acute thalamic ischemic stroke," Comput. Math. Methods Med. **2016**, e2582478.
[78]R. K. O. Chu, A. R. Braun, and J. A. Meltzer, "MEG-based detection and localization of perilesional dysfunction in chronic stroke," NeuroImage Clin. **8**, 157–169 (2015).
[79]R. Sun, W.-W. Wong, J. Gao, G. F. Wong, and R. K.-Y. Tong, "Abnormal EEG complexity and alpha oscillation of resting state in chronic stroke patients," in *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)* (IEEE, 2021), pp. 6053–6057.
[80]A. Delgado-Bonal and A. Marshak, "Approximate entropy and sample entropy: A comprehensive tutorial," Entropy **21**, 541 (2019).
[81]A. Omidvarnia, M. Mesbah, M. Pedersen, and G. Jackson, "Range entropy: A bridge between signal complexity and self-similarity," Entropy **20**, 962 (2018).
[82]J. A. Palmer, L. A. Wheaton, W. A. Gray, M. A. Saltão da Silva, S. L. Wolf, and M. R. Borich, "Role of interhemispheric cortical interactions in poststroke motor function," Neurorehabil. Neural Repair **33**, 762–774 (2019).
[83]E. P. Casula, M. C. Pellicciari, S. Bonnì, B. Spanò, V. Ponzo, I. Salsano, G. Giulietti, A. Martino Cinnera, M. Maiella, I. Borghi *et al.*, "Evidence for interhemispheric imbalance in stroke patients as revealed by combining transcranial magnetic stimulation and electroencephalography," Hum. Brain Mapp. **42**, 1343–1358 (2021).
[84]T. Paul, V. M. Wiemer, L. Hensel, M. Cieslak, C. Tscherpel, C. Grefkes, S. T. Grafton, G. R. Fink, and L. J. Volz, "Interhemispheric structural connectivity underlies motor recovery after stroke," Ann. Neurol. **94**, 785–797 (2023).
[85]N. Meneghetti, C. Cerri, E. Vannini, E. Tantillo, A. Tottene, D. Pietrobon, M. Caleo, and A. Mazzoni, "Synaptic alterations in visual cortex reshape contrast-dependent gamma oscillations and inhibition-excitation ratio in a genetic mouse model of migraine," J. Headache Pain **23**, 125 (2022).
[86]N. Meneghetti, C. Cerri, E. Tantillo, E. Vannini, M. Caleo, and A. Mazzoni, "Narrow and broad γ bands process complementary visual information in mouse primary visual cortex," eNeuro **8**, ENEURO.0106-21.2021 (2021).
[87]C. Magri, K. Whittingstall, V. Singh, N. K. Logothetis, and S. Panzeri, "A toolbox for the fast information analysis of multiple-site LFP, EEG and spike train recordings," BMC Neurosci. **10**, 81 (2009).
[88]S. Panzeri, R. Senatore, M. A. Montemurro, and R. S. Petersen, "Correcting for the sampling bias problem in spike train information measures," J. Neurophysiol. **98**, 1064–1072 (2007).
[89]C. Ding and H. Peng, "Minimum redundancy feature selection from microarray gene expression data," in *Computational Systems Bioinformatics. CSB2003. Proceedings of the 2003 IEEE Bioinformatics Conference. CSB2003* (IEEE, 2003). pp. 523–528.
[90]GitHub, see https://github.com/nicolomeneghetti/Recovery_Stroke_Prediction.