Brain fingerprint is based on the aperiodic, scale-free, neuronal activity

Pierpaolo Sorrentino, Emahnuel Troisi Lopez, Antonella Romano, Carmine Granata, Marie Constance Corsi, Giuseppe Sorrentino, Viktor Jirsa

NeuroImage 277 (2023) 120260

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Brain fingerprint is based on the aperiodic, scale-free, neuronal activity

Pierpaolo Sorrentino [a,b,c,1,*], Emahnuel Troisi Lopez [c,d,1], Antonella Romano [d], Carmine Granata [c], Marie Constance Corsi [e], Giuseppe Sorrentino [c,d,f], Viktor Jirsa [a]

[a] Institut de Neurosciences des Systèmes, Aix-Marseille Universitè, Marseille, France
[b] Deparment of Biomedical Science, University of Sassari, Sassari, Italy
[c] Institute of Applied Sciences and Intelligent Systems, CNR, Naples, Italy
[d] Department of Motor Sciences and Wellness, University of Naples "Parthenope", Naples, Italy
[e] Sorbonne Université, Institut du Cerveau - Paris Brain Institute - ICM, CNRS, Inria, Inserm, AP-HP, Hôpital de la Pitié Salpêtrière, F-75013, Paris, France
[f] Institute of Diagnosis and Treatment Hermitage Capodimonte, Naples, Italy

ARTICLE INFO

Keywords:
Neuronal Avalanches
Brain Dynamics
Brain Differentiability
Transition Matrices
Magnetoencephalography

ABSTRACT

Subject differentiation bears the possibility to individualize brain analyses. However, the nature of the processes generating subject-specific features remains unknown. Most of the current literature uses techniques that assume stationarity (e.g., Pearson's correlation), which might fail to capture the non-linear nature of brain activity. We hypothesize that non-linear perturbations (defined as neuronal avalanches in the context of critical dynamics) spread across the brain and carry subject-specific information, contributing the most to differentiability. To test this hypothesis, we compute the avalanche transition matrix (ATM) from source-reconstructed magnetoencephalographic data, as to characterize subject-specific fast dynamics. We perform differentiability analysis based on the ATMs, and compare the performance to that obtained using Pearson's correlation (which assumes stationarity). We demonstrate that selecting the moments and places where neuronal avalanches spread improves differentiation ($P < 0.0001$, permutation testing), despite the fact that most of the data (i.e., the linear part) are discarded. Our results show that the non-linear part of the brain signals carries most of the subject-specific information, thereby clarifying the nature of the processes that underlie individual differentiation. Borrowing from statistical mechanics, we provide a principled way to link emergent large-scale personalized activations to non-observable, microscopic processes.

1. Introduction

Higher cognitive functions rely on the complex, coordinated interactions among multiple brain areas. Whole-brain functional measurements (e.g., fMRI, M/EEG) have been used to characterize such interactions, typically using the pair-wise statistical dependencies between regional signals as a proxy (Bullmore and Sporns, 2009). The time-averaged patterns of interactions (estimated via the Pearson's or Spearman's correlation coefficient in fMRI or via synchronization metrics (among others) in M/EEG) form the static functional connectome (sFC) (Friston, 2011). sFCs contain subject-specific information, constituting a "brain fingerprint" which allows individual differentiation (Amico and Goñi, 2018; da Silva Castanheira et al., 2021; Finn and Rosenberg, 2021).

However, large-scale brain activity reconfigures itself over time (Zalesky et al., 2014). A recent fMRI-based study used a sliding-window approach to show that specific time intervals carry the vast majority of the subject-specific information (Van De Ville et al., 2021). With respect to MEG data, it has been shown that data segments as short as 30 seconds contain enough information to unambiguously differentiate individuals (da Silva Castanheira et al., 2021). These findings pose a challenge on the nature of the information upon which subject differentiation is based. In fact, the intermittent, short-lived nature of the mo-

Abbreviations: AAL, Automated Anatomical Labeling; ATM, Avalanche Transitions Matrix; DM, differentiation matrix; EOG, electro-oculogram; FC, Functional Connectome; ICA, independent component analysis; MEG, Magnetoencephalography; LMCV, Linearly Constrained Minimum Variance; PC, Pearson's Correlation Coefficient; PCA, principal component analysis; ROIs, Regions of Interest; SC, Spearman's Correlation Coefficient; sFC, static Functional Connectome; SR, Success Rate.

* Correspondence to: Dr. Pierpaolo Sorrentino, Institut de Neurosciences des Systèmes, Aix-Marseille Universitè, 27, Boulevard Jean Moulin, 13005 Marseille, France

E-mail address: pierpaolo.sorrentino@univ-amu.fr (P. Sorrentino).

[1] The authors contributed equally.

https://doi.org/10.1016/j.neuroimage.2023.120260.
Received 1 March 2023; Received in revised form 13 June 2023; Accepted 28 June 2023
Available online 29 June 2023.

ments of "high differentiability" might be generated by non-stationary processes (Deco et al., 2011).

Converging evidence shows that large-scale brain activity has an aperiodic, bursty component, which mainly drives static connectivity (Zamani Esfahlani et al., 2020). The presence of large bursts is not expected for stationary processes, whereas it can be a manifestation of a dynamical system operating in a near-critical regime (characterized by a branching ratio $\sim= 1$), in the sense of diverging correlation lengths (Chialvo, 2010). In particular, while recent evidence showed that high excursions can occur in the case of a stationary process (Novelli and Razi, 2022), the features of the bursty activity as observed in MEG are best explained by not-assuming stationarity (Sorrentino et al., 2023). Borrowing from statistical mechanics, these scale-free bursts have been conceptualized as "neuronal avalanches" (Shriki et al., 2013a). Neuronal avalanches are fine-tuned, present across species, across temporal and spatial scales, and have been observed with multiple techniques (Beggs and Plenz, 2003). In the human brain, neuronal avalanches spread across gray matter regions through the white-matter bundles linking them (Sorrentino et al., 2021c). Furthermore, altered avalanche dynamics is related to clinical disability in multiple neurological diseases (Polverino et al., 2022; Sorrentino et al., 2021a), demonstrating the clinical relevance of this phenomenon.

Here, we hypothesize that neuronal avalanches are the expression of subject-specific, large-scale brain processes. As a consequence, we expect individual differentiation based on avalanches to perform better as compared to using the entire available data (for instance, using the sFC). To test our hypothesis, we used data from a previously published cohort (Sorrentino et al., 2021c), consisting of source-reconstructed MEG data from 44 healthy subjects. Each participant underwent two magnetoencephalography (MEG) scans, separated by a ~1.5 minute-long pause. We then confirmed our data in two additional datasets, where the two MEG scans were separated by 12 hours and by one month, respectively (see supplementary material 1). To identify avalanches, each source-reconstructed signal was z-scored (across-time) and then set to 1 if above a threshold, and to 0 otherwise. Each avalanche was defined as starting when at least one region was above threshold, and as finishing when no region showed unusually high activation. Note that the higher the threshold, the less data is considered for differentiation. For each avalanche, the ij^{th} entry of the transition matrix contains the probability of region j being recruited at time t+1, given that region i had been recruited at time t. We then averaged, edgewise, within each subject, all the avalanche-specific transition matrices and we obtained, per each session, a subject-specific avalanche transition matrix (ATM), which captures the spatio-temporal dynamics of neuronal avalanches in each individual. Then, we used the correlations between the ATMs as a metric of similarity (between the first and the second MEG scans), and used it to define individual differentiability (i.e., if two ATMs from one subject resemble each other more than they resemble the ATMs from other participants, the subject is correctly differentiated). After comparing each individual with all the others, we averaged the results (per subject) obtaining a score defined as success rate (SR), which refers to the ability in differentiating an individual in that specific dataset. See methods for further details.

First, we set out to define the optimal threshold to maximize differentiability. Under the hypothesis that the spreading of aperiodic, scale-free perturbations carries subject-specific information, we expect maximal differentiation to occur when focusing on the scale-free activity, as opposed to taking the whole data into account. Hence, we varied the threshold used to define a region as active. We hypothesize the SR to peak when linear activity is discarded and scale-free activity is preserved.

Then, we compared the results obtained with our approach (i.e., the ATM) with three alternative approaches (Fig. 1A). First, we performed differentiation analysis based on connectivity defined as the pairwise Spearman's correlation coefficients (SC) on the whole dataset (FC full dataset). The same procedure was also repeated using pairwise Mutual Information (see supplementary materials 2). Given the hypothesis that information-carrying interactions occur on the large-scale during neural avalanches, we reasoned that differentiation should be worse using SC as compared to ATM, despite the fact that SC exploits the whole data while the ATMs are only based on a few locations and time points (i.e., the specific moments when an avalanche has recruited specific regions). Then, we selected the moments when neuronal avalanches were occurring, and concatenated the corresponding z-scored time series (i.e., before binarization) (FC when avalanches). Based on this (small) part of the original data, we performed differentiation analysis. In this case, we expect worse differentiability as compared to the ATM, but comparable to the SC based on the whole signal. Finally, we have repeated the same procedure, this time by randomly selecting intervals of the z-scored time series where no avalanche was occurring, and concatenated them to obtain avalanche-free data of the same length as the total length of all the avalanches (FC when no avalanches). We expect the lowest differentiation in this case (in fact, according to our hypothesis, we would have excluded the most informative parts of the signal from our analysis). We repeated this procedure 1000 times, hence providing a null-distribution representing the SRs to be expected when selecting random segments of the data, under the null hypothesis that avalanches do not carry most of the subject-specific information. An overview of the procedure is provided in Fig. 1B.

2. Methods

2.1. Cohort, data acquisition, preprocessing, source-reconstruction, avalanches estimation

Methods for the description of the cohorts and the data acquisition, preprocessing, source-reconstruction, avalanches estimation and avalanche transition matrices have been described in detail in Liparoti et al. (2021), Pesoli et al. (2022), and Sorrentino et al. (2021c). In short, we recruited 58 young adults (male 32 / female 26, mean age ± SD was 30.72 ± 11.58), right-handed and native Italian speakers with no major internal, neurological or psychiatric illnesses and no use of drugs or medication that could interfere with MEG/MRI signals. The study complied with the Declaration of Helsinki and was approved by the local Ethics Committee. All participants gave written informed consent. 3D T1-weighted brain volumes were acquired at 1.5 Tesla (Signa, GE Healthcare) after the MEG recording (Sorrentino et al., 2021c). The MEG registration was divided in two eyes-closed segments of 3:30 minutes each, separated by a ~1.5-minute-long break.

For supplementary analyses (supplementary materials 1), two more datasets including MEG recordings that took place several hours/days apart were taken into consideration. Dataset 1 included twenty-three male participants (mean age ± SD, 25.14 ± 2.53 years), who performed the first session of recording at 9 a.m., and the second session at 9 p.m. of the same day. Dataset 2 included twenty-six female participants (mean age ± SD, 26.6 ± 5.1 years). In this case, the first and the second MEG sessions took place 28 ± 1.5 days apart. Acquisition methods, preprocessing pipelines and data analysis were the same for all three datasets.

To identify the position of the head, four anatomical points and four position coils were digitized. Electrocardiogram (ECG) and electrooculogram (EOG) signals were also recorded. After an anti-aliasing filter, MEG signals were acquired at 1024 Hz, then a fourth order Butterworth IIR band-pass filter in the 0.5-48 Hz was applied. Noisy channels were identified and removed manually, Principal Component Analysis (PCA) and supervised Independent Component Analysis (ICA) were used to remove environmental noise and physiological artifacts from the ECG (one component) and the EOG (zero component), respectively. The whole cleaning procedure was performed by an expert rater, who excluded 14 participants, as the signals were too noisy. Hence, 44 subjects were selected for further analysis. The time series of neuronal activity were reconstructed in 116 regions of interests (ROIs), based on the Automated Anatomical Labeling (AAL) atlas, using the Linearly Constrained Min-

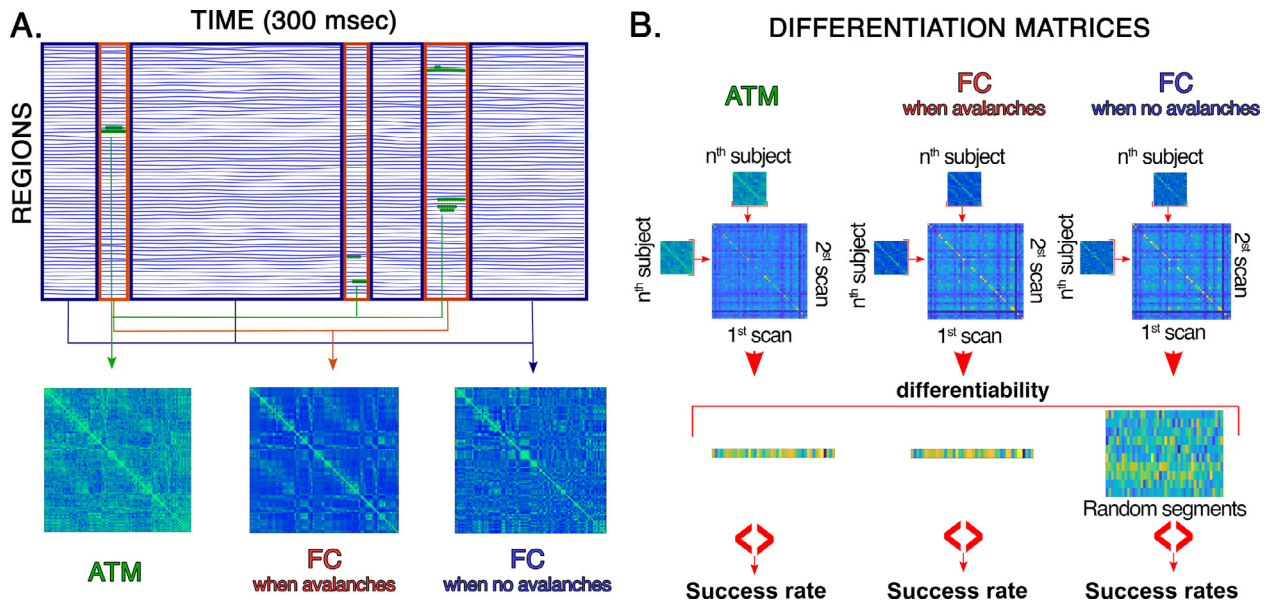

Fig. 1. Pipeline overview (A) Data selection. Blue lines represent z-scored source-reconstructed time-series. In green, the data points that are above threshold. The transitions between these points are the base to compute the ATM. The segments of the data when avalanches are occurring are surrounded by red rectangles. All these segments are concatenated, and the Spearman's correlation is computed in the "FC when avalanches" case. Segments without avalanches are surrounded by blue rectangles. These constitute the vast majority of the data. These segments are concatenated multiple times, each time with a random selection of them, until the same length of the "avalanche" segment is reached ("FC when no avalanches"). The pipeline analysis also includes FC matrix computed on the whole recording ("FC full data"), that is not shown in the figure. **(B)** Construction of the differentiation matrices, containing the Spearman's correlation coefficient between the connectivity matrix of the first and the second acquisitions. For subject *i*, the success rate is defined as the number of participants that are less similar to subject *i* as compared to subject *i* itself. The success rate refers to the average of the success rates across subjects.

imum Variance (LCMV) beamformer algorithm, and the native structural MRIs. The sources were reconstructed for the centroids of each ROI (Hillebrand et al., 2016). Finally, we considered a total of 90 ROIs for the AAL atlas, since we excluded the cerebellum because of its lower reliability in MEG (Andersen et al., 2020). All the preprocessing steps and the source reconstruction were made using the Fieldtrip toolbox (Oostenveld et al., 2010). To study the dynamics of brain activity, we estimated "neuronal avalanches" from the source-reconstructed time series. Firstly, the time series of each ROI was discretized calculating the z-score, as in:

$$X(t) = \frac{x(t) - \mu}{\sigma} \tag{1}$$

where x is the signal, μ is the mean of the signal over time, and σ is its standard deviation.

Then, we identified positive and negative excursions beyond a threshold, as:

$$X_t(t) = \begin{cases} 1 & if\ abs\left(X_i(t)\right) > Thres \\ 0 & otherwise \end{cases} \tag{2}$$

The value of the threshold was varied between 1.5 and 3.5, to spot the value that maximized differentiation. A neuronal avalanche begins when, in a sequence of contiguous time bins, at least one ROI is active (i.e., above threshold), and ends when all ROIs are inactive (Sorrentino et al., 2021a). The branching parameter was geometrically averaged across avalanches and subjects. In fact, systems operating at criticality typically display a branching ratio ~1. In detail, the branching ratio was calculated as the geometrically averaged (over all the time bins) ratio of the number of events (activations) between the subsequent time bin (descendants) and that in the current time bin (ancestors) and then averaging it over all the avalanches, as:

$$\sigma_i = \prod_{j=1}^{N_{bin}-1} \left(\frac{n_{events}\ (j+1)}{n_{events}\ (j)} \right)^{\frac{1}{N_{bin}-1}} \tag{3}$$

$$\sigma = \prod_{i=1}^{N_{aval}} \left(\sigma_i\right)^{\frac{1}{N_{aval}}} \tag{4}$$

Where σ_i is the branching parameter of the i-th avalanche in the subject, N_{bin} is the total amount of bins in the i-th avalanche, N_{aval} is the total number of avalanches in the dataset. To allow a better characterization of the statistical properties of the avalanches, we also report in Supplementary materials 3 and 4, the number of avalanches at each threshold, the exponents of the distribution of the sizes of the avalanches (where the size is defined as the number of active regions), and that of the durations. The results shown are derived when considering all avalanches. However, we repeated the analysis taking into account only avalanches longer than 5 time bins, as well as only avalanches longer than 20 time bins (this further reduced the selected datapoints), and the results were unchanged (see supplementary material 5). Then, for each avalanche n, the transition matrix AvalATM (n) was defined as:

$$AvalATM_{i,j}(n) = P(X_j(t + \delta t) > Thres \mid X_j(t) > Thres) \tag{5}$$

where the element (*i, j*) represents the probability that region *j* is active at time *t*+δ, given that region *i* was active at time *t*, where δ~3ms. The ATMs were averaged per participant (i.e., across avalanches), as

$$ATM_{ij} = \frac{1}{Naval} \sum_{n=1}^{Naval} AvalATM_{ij}(n) \tag{6}$$

, and finally symmetrized. The introduction of a time-lag makes it unlikely that our results can be explained trivially by field spread (i.e., the fact that multiple sources are detected simultaneously by multiple sensors, generating spurious zero-lags correlations in the recorded signals). Please also refer to (Sorrentino et al., 2021c), for extensive analysis on field spread on ATMs in this dataset. For comparison, besides the ATM, also sFC was computed based on both Pearson's and Spearman's correlation coefficient.

2.2. Fingerprint analysis

Fingerprinting analysis was performed similarly to Romano et al. (2022), Sorrentino et al. (2021b), and Troisi Lopez et al. (2023). In short, for each subject, for each of the two acquisitions, an adjacency matrix was built (either for the ATM or the ones based on the sFC). Hence, each participant had two matrices, each referring to one of the two acquisitions. Hence, a differentiation matrix was built (Fig. 1, panel B), with rows and columns referring to the first and the second acquisitions for each participant, respectively. The differentiation matrix (DM) was defined, based on the Pearson's correlations between the adjacency matrices (AM, which in turn were based on either the ATMs or on sFC) derived from the first and the second scans, as:

$$DM(a_1,b_2)=\frac{\sum\left(tril\left(AM(a_1)\right)-\left\langle tril\left(AM(a_1)\right)\right\rangle\right)\left(tril\left(AM(b_2)\right)-<tril(AM(b_2))>\right)}{\sqrt{\sum\left(tril(AM(a_1))-\left\langle tril\left(AM(a_1)\right)\right\rangle\right)^2\sum\left(tril\left(AM(b_2)\right)-\left\langle tril\left(AM(b_2)\right)\right\rangle\right)^2}} \quad (7)$$

Where a_1 denotes the first scan for subject a, and b_2 indicates the second scan for subject b, "tril" extracts the lower triangular part of the adjacency matrix, and "< >" indicates the average. Furthermore, we used the Mantel test to demonstrate the robustness of the correlations (Glerean, 2014).

Hence, the main diagonal contains the similarity between each subject with themselves between the first and the second acquisition. The off-diagonal elements contain the similarities of each subject with every other subject. Hence, if the value of a subject on the main diagonal is the highest, that subject is correctly differentiated. To obtain the individual success rate, referring to how much an individual is differentiated, we first calculated the number of times that the similarity between two AMs of the same individual was higher than the similarity between the AMs of that given individual with the ones of each other subjects; then we divided the results by the number of individuals in the dataset, as in:

$$SR_a=\frac{1}{N\ Subj-1}\left[N\ Subj-R\left(DM_{a,a}\right)\right] \quad (8)$$

where R indicates the rank of the elements in position (a,a) along column a. The average of the individual success rates was defined as the success rate (SR), which refers to the whole population, and was used as a readout in our analysis, defined as:

$$SR=\frac{1}{N\ Subj}\sum_{a=1}^{N\ Subj}SR(a) \quad (9)$$

Furthermore, the 95% confidence intervals were calculated using the bootstrapping method. Specifically, we performed 1000 random subsampling equal to 90% of the full sample (40 out of 44) and calculated the 95% confidence intervals within the range 2.5 - 97.5 percentiles.

Finally, to assess the effect size and to compare the performance of the ATM approach with respect to the sFCs approach, we calculated, for each method, the ratio between the SR and the amount of information used to calculate the respective SR. Hence, we computed two parameters SR/time and SR/bit. The former divides the SR by the amount of time (in seconds) required by each approach. The latter divides the SR by the amount of data (bit) required by each approach, calculated by multiplying the number of recording frames (temporal component) for the number of regions considered (spatial component). It must be noticed that while sFCs approaches use the whole connectome, hence 90 brain regions per frame, ATMs only include the active regions at each different frame. The two parameters were compared between the ATM and the three sFCs using the Wilcoxon signed rank test, and the p-values were corrected by false discovery rate (Benjamini & Hochberg, 1995).

Finally, the resiliency to noise of the ATMs was compared to that of the FCs in Supplementary material 6.

3. Results

In this paper, we set out to quantify individual differentiability using the success rate (SR), defined as the average individual success rates (which captures the extent to which an individual is differentiated within a specific dataset). We have explored the SR as a function of the threshold used to binarize the time-series. This was done under the hypothesis that avalanches define the most "informative" moments with respect to subject-specific activities. If this were not the case, differentiation should drop as the threshold grows, provided that a higher threshold discards more data (note that a z-score of 2.8 discards ~99% of a z-score distribution). Our results show that the maximal SR is obtained with a threshold equal to $z > |2.8|$ (Fig 2A, green line). The differentiation rate improves as the threshold grows from $|1.5|$ to $|2.8|$, despite the fact that less data is being taken into account, and this in accordance with the hypothesis that subject differentiation is based on the spatio-temporal dynamics of neuronal avalanches (which would be, in turn, the main driver of functional connectivity).

Then, for each threshold, we selected from the original z-scored time series the timepoints when avalanches had been occurring (see Fig. 1, panel A). Based on these time points, we have computed the adjacency matrices using the sFC and used these, instead of the ATMs, to perform subject differentiation. In other words, we are now using data from the time interval in which avalanches were occurring, but we are not selecting the specific regions that were recruited by each avalanche (as we consider all the regions instead). Using this approach, we show that the SR is lower using the sFC as compared to using the ATMs (Fig. 2, panel A, the red line indicates, for each threshold, the SR based on the sFC, while the green line denotes the SR obtained using the ATMs). This might be compatible with the idea that large-scale interactions occur at intermittent time-points and in specific brain regions.

Then, for each threshold, we have randomly selected a number of timepoints equal to the lengths of all the avalanches occurring in that given recording, but this time choosing among the moments when no avalanche was occurring. Based on this choice, we have again observed the individual differentiation. We show that, despite the fact that we consider the same amount of data, the differentiation is worse if neuronal avalanches are excluded from the analysis (Fig. 2A, blue line). Finally, even the differentiation success rate based on the full recordings (Fig. 2A, yellow line, which is independent from the thresholds), displays a lower value with respect to the ATMs. For each approach, 95% confidence intervals within the 2.5 - 97.5 range were computed (see supplementary materials 7), and the success rate of the ATMs was above the upper confidence interval of the analyses based on sFCs.

In Fig. 2B, we report the SR based on ATM (green), on sFC based on avalanches (red), on sFC based on the whole data (yellow), and on the 1000 instances of the sFC computed in moments when no avalanches were occurring (blue distribution), using a threshold equal to $z > |2.8|$. This shows that, using sFC, the best estimates are obtained from the moments when avalanches are occurring, and this is unlikely to be obtained given a random selection ($P < 0.001$). In Fig. 2C, we reported the similarities between the ATMs based on the first (rows) and second (columns) acquisitions. Higher values on the main diagonal denote successful differentiation. The number of non-differentiated individuals (i.e., the subjects that are not differentiated against all the others) corresponds to sixteen for ATM, FC when avalanches, FC when no avalanches, while it corresponds to seventeen for FC on the whole recording. However, the fractional differentiation shows that even when an individual is not completely differentiated, ATMs outperform the other approaches, even though they only use both very short time windows and reduced spatial information (i.e., only the region recruited in each avalanche are included in the computation of the ATMs). Furthermore, the Mantel test performed on the differentiation matrices confirmed 99% of the significance values.

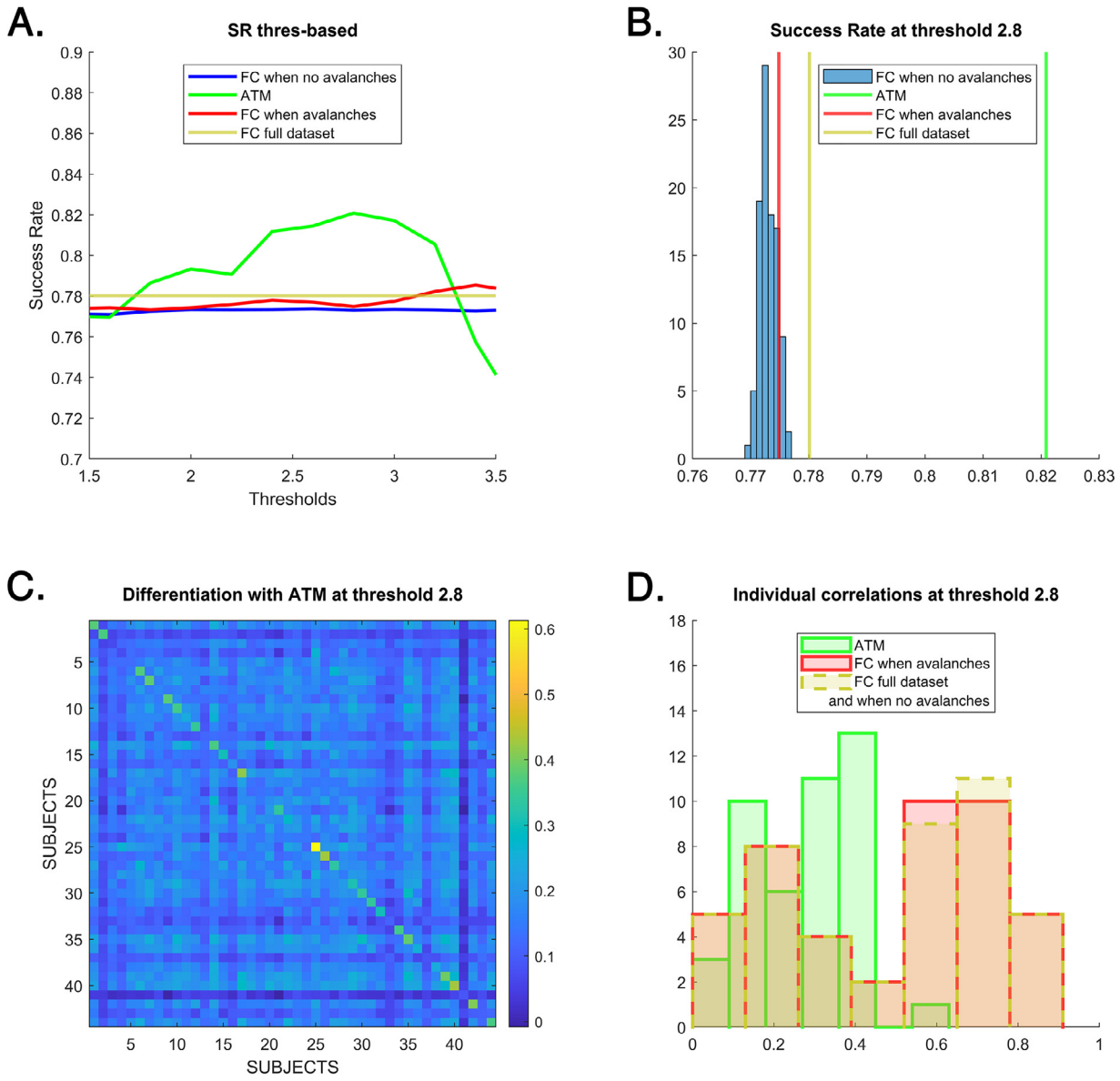

Fig. 2. Differentiation analysis (A) Success rate as a function of the threshold defining avalanches. The green, red, blue, and yellow lines refer to the ATM, the FC when there are avalanches, the FC in moments when no avalanches are present, and FC on the whole recording. **(B)** for the threshold set to z=|2.8| (i.e., the threshold corresponding to the best success rate), the green, red, and yellow lines correspond to the SR obtained using ATM, FC computed only in moments when there were avalanches, FC computed on the whole dataset, respectively. The distribution in blue refers to 1000 SRs obtained for 1000 random selections of moments when no avalanche was present, each one of the same length of the moments with avalanches. **(C)** Differentiation matrix, containing the Spearman's correlation coefficient between the matrices belonging to first and second acquisitions. The main diagonal contains the similarities of each subject with itself. **(D)** The histograms and corresponding distributions refer to the values of the correlation between the first and second acquisitions for each individual. The green, red, and yellow distributions refer to the similarities between the ATMs, the FC based on moments when avalanches are present, and the FC on the whole recording (that graphically overlaps the FC when no avalanche is present), respectively.

Then, we moved on to compare the similarities (correlation) between the adjacency matrices based on the sFCs (when avalanches were occurring) and those based on the ATMs. Fig. 2D shows the distribution of the correlations for the ATMs (green), for the sFCs computed on the whole data (yellow, dotted edges), and for the sFCs computed on the moments when avalanches were occurring (red). The distribution of FC when avalanches were not occurring overlaps with the distribution of sFCs computed on the whole data. It is easy to see that the ATMs show lower correlation values, as compared to sFC. The Kolmogorov Smirnov test confirmed that the ATMs differ from the sFCs ($p < 0{,}0001$), which in turn are not statistically different between themselves ($p = 0.98$). However, as shown, higher auto-correlations do not translate into better differentiability, as this feature also depends on the similarity between different individuals (i.e., the correlations with other individuals). In this case, the results suggest that high auto-correlation may be not driven by subject-specific information. Specifically, the higher auto-correlation values (obtained with FC) may be determined by features that are shared among participants. Hence, despite higher correlation values, subject differentiation does not improve. Finally, the SR based on the mutual information computed on the whole data is equal to 0.76 (supplementary material 2).

Furthermore, as estimation of the effect size of the differentiation outcomes, we calculated the ratio between the SR and the amount of information used for each approach/metric. Fig. 3A, shows the SR/time difference between ATMs and the sFCs analyses. ATMs SR/time resulted to be significantly higher than sFCs when avalanches ($z = 2.3$,

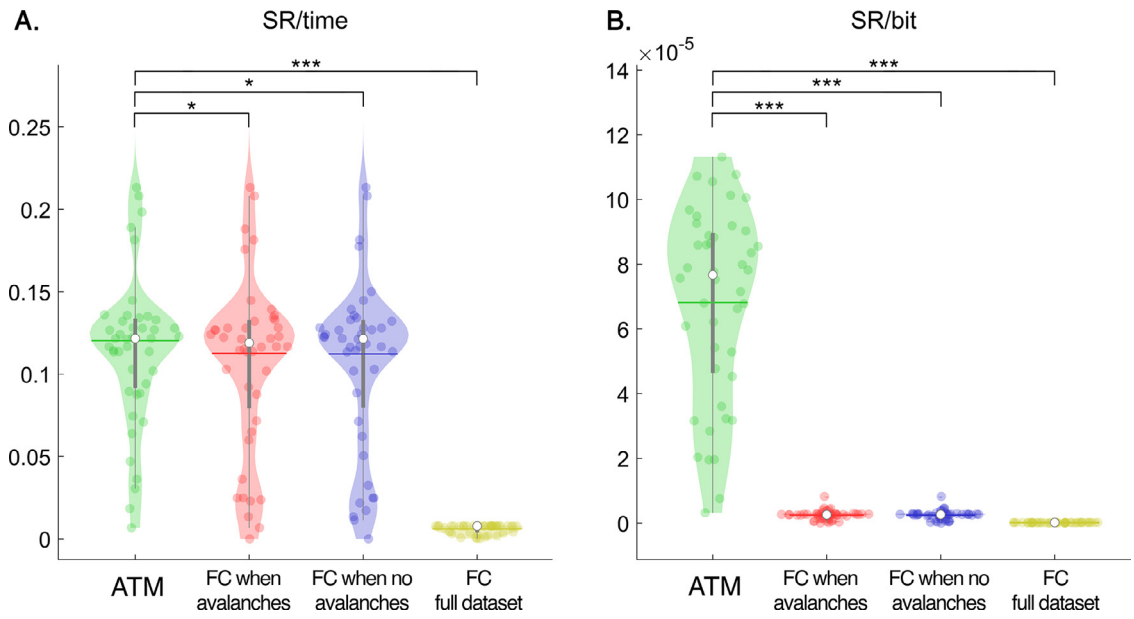

Fig. 3. Performance comparison. Violin plot of the comparison of ATM performance vs FC approaches. (A) Comparison of the SR/time parameter, that corresponds to the successful differentiation divided by the amount of time (in seconds) considered for each given approach. (B) Comparison of the SR/bit parameter, that corresponds to the successful differentiation divided by the amount of data (i.e. the frames considered in each approach multiplied by brain region included in each analysis). * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ after false discovery rate correction.

pFDR = 0.026), sFCs when no avalanches (z = 2.02, pFDR = 0.043), and sFCs total (z = 5.78, pFDR < 0.001). Furthermore, Fig. 3B displays the comparison when also the spatial component is kept into consideration. ATMs SR/bit resulted to be significantly higher than sFCs when avalanches (z = 5.78, pFDR < 0.001), sFCs when no avalanches (z = 5.78, pFDR < 0.001), and sFCs total (z = 5.78, pFDR < 0.001)

Success rate of differentiation was also analyzed using both supplementary datasets 1 (two recordings 12 occurred hours apart) and dataset 2 (two recordings occurred ~28 days apart). With regard to the dataset 1, the success rates were 0.675 for the ATMs, 0.67 for FC full dataset, 0.662 for FC when avalanches, and peak equal to 0.664 for the FC when no aval distribution. Concerning dataset 2, the success rates were 0702 with ATM, 0.682 for FC full dataset, 0.66 for FC when avalanches, and a peak equal to 0.66 FC when no avalanches. Again, the success rates obtained with ATMs were higher than the ones based on functional connectivity (see supplementary material 1).

4. Discussion

In this manuscript, we test the role of aperiodic, scale-free, higher-order activations related to brain dynamics (i.e., neuronal avalanches) in individual differentiation. Avalanches are defined as occurring when regions deviate from their baseline activity, which is defined, in turn, on the standardized score of the regional signal. In other words, we select the moments and locations (regions) which show unexpected levels of activity (given the baseline, linear activity of each region), thereby focusing on higher-order activity. In particular, we verified that neuronal avalanches were characterized by a branching ratio = 1, which is a hallmark of a branching process operating at criticality. The analyses of the exponents of the sizes and duration of the avalanches are reported in the supplementary materials 4. In fact, neuronal avalanches are well-defined objects, whose properties have been predicted by mean-field theory, and empirically verified in human data (Shriki et al., 2013b). Our results appear in line with the literature about avalanches in MEG data in humans.

Then, we use the recently described avalanche transition matrix (ATM) to capture the spatio-temporal structure of avalanches (Sorrentino et al., 2021c). The avalanche-specific transition matrix contains, in the ij^{th} position, the probability that region j will show unusual high activity after region i did. In other words, the transition matrix contains the probability that a "wave" of higher-order activations would propagate from i to j. Averaging across avalanche-specific transition matrices, we obtained one subject- and session- specific ATM and utilized the similarity between session-specific ATMs to perform individual differentiation. We show that fingerprint analysis performed on ATMs improves performance as compared to using matrices based on the correlation coefficient, despite the fact that significantly more data is considered in the latter case. We have replicated the results using Pearson's correlation instead of Spearman's, and the results are confirmed (not shown). In fact, by construction, avalanches only select the rare, fat-tailed part of the regional activity, and entirely discard the vast majority of data points. However, these few, scattered moments and locations are the ones carrying most subject-specific information. Indeed, we assessed the informativeness of the time-points taken into account (SR/time and SR/bit), which highlighted the relevance of the avalanches toward differentiation, as compared to techniques that do not select datapoint in space and time.This might talk to the converging evidence showing that aperiodic, scale-free dynamics conveys physiologically meaningful and subject-specific activity (Beggs and Plenz, 2003; Rucco et al., 2020; Shriki et al., 2013a; Sorrentino et al., 2021a; Zamani Esfahlani et al., 2020). Accordingly, the spreading of neuronal avalanches, at the individual level, is related to the individual structure of white-matter bundles (Sorrentino et al., 2021c). This suggests the idea that an input received by a region (from the rest of the brain) might affect the local processes, provoking an "unusual" activation (Rabuffo et al., 2021). While purely speculative, one could frame this within the "communication through coherence" hypothesis, whereby the incoming inputs would "entrain" local neuronal populations, momentarily favoring their synchronization, and resulting in a signal amplitude that would have not been generated in the absence of the incoming signal (Fries, 2015). As said, once such perturbation is generated, it spreads preferentially across the white-matter bundles. Recently, a mechanistic, mean-field based model of the brain dynamics showed that avalanche-like activity might be generated when regions are realistically coupled (Porta and Copelli, 2019). In conclusion, our data show that ATMs provide a mathematically and physiologically

rooted, yet straight-forward way to describe subject-specific fast dynamics. However, it is important to stress that our analysis is not comprehensive in terms of aperiodic activities in the brain. In particular, recent works focused specifically on the aperiodic part of the activities as captured by the 1/f slope of the power-spectra, which is also relevant to behaviour and related to the biological features of individuals (e.g., to the age) (Cellier et al., 2021; Donoghue et al., 2020). In fact, evidence show that these features also convey individual features, in health and diseased population (da Silva Castanheira et al., 2021), and allow subject differentiation (Demuru and Fraschini, 2020). However, the relationship between neuronal avalanches and the 1/f activities remains an open question that will require further investigations.

The correlation values between the sFC-based adjacency matrix from the two sessions were generally significantly higher than those obtained using the ATMs. This is somewhat unsurprising, considering that the former exploits the whole dataset, while the latter limits itself to a few data points. However, the sFCs are outperformed by the ATMs in terms of individual differentiation. Hence, one might interpret that the sFCs capture features that are shared by the whole dataset, while the ATMs focus more specifically on subject-specific dynamics. Furthermore, the ATMs do not consider zero-lag associations, hence correcting for field spread, while SC is biased in that sense (Sorrentino et al., 2021c). Connectivity metrics that do not correct for signal leakage tend to outperform, in terms of fingerprinting, those who do. This is interpreted as an effect of the contribution of signal leakage to differentiation, provided that the geometry of the head shapes the field-spread, and that this carries subject-specific information, although not neurophysiological in nature (Demuru et al., 2017). In this sense, the higher similarities obtained by the SC might not be due to similar patterns of activity but, rather, by similar structural features. In conclusion, ATMs are effective in selecting the locations and moments where higher-order perturbations spread, which might be directly related to subject-specific neurophysiological mechanisms. In fact, when sFCs is performed on the data segments that contain the neuronal avalanches, the performance remains fairly high, confirming that the avalanches indeed represent the moments where subject-specific dynamics is manifesting itself. Accordingly, when subject differentiation is analyzed based on moments when the avalanches are not occurring, the most informative part of the signal is lost, and the performance never reaches the levels observed when avalanches are included. Finally, we show in supplementary materials 6 that subject-differentiation based on the ATMs is resilient against high levels of noise.

The fact that the results are confirmed in the supplementary datasets, in which recordings are carried out several hours/days apart, confirmed that it is unlikely that the results observed in the main analysis driven by environmental noise or, in general, from the conditions in which the acquisitions took place. Finally, our results will need further validation in other datasets. In particular, the cohort we present here is middle-sized, as compared to publicly available datasets. Further studies should also investigate the possible relationship between ATMs characteristics and demographic or behavioral aspects of the sample. Finally, when interpreting the results, a possible limitation may be related to the choice of using the AAL atlas. Indeed, one should take into account that the AAL has been developed using volumetric data from a single subject, which might lead to errors when applied to more subjects (Tzourio-Mazoyer et al., 2002). Therefore, the individual differentiability may be biased. On the other hand, one needs to consider that the AAL and the MEG have comparable resolutions. Further studies should investigate the differentiability performance using different brain parcellations (Eickhoff et al., 2018).

In conclusion, the ATMs appear as a straight-forward, yet principled way to select the data that convey subject-specific, large-scale spatio-temporal dynamics. This opens new venues to characterize subject-specific brain dynamics in health and disease, and provides new observables to tune subject-specific brain models. Furthermore, it corroborates the idea that large-scale interactions are of higher order (i.e., non-expected by a linear process), and should be treated as such when processing brain signals. The ATM is a useful tool to quantify such higher-order large-scale dynamics. Finally, task-based studies might be analyzed using ATMs, in order to identify the role of specific regions and/or edges in specific behavioral functions.

Funding

This work was supported by the European Union's Horizon 2020 research and innovation program under grant agreement No. 945539 (SGA3) Human Brain Project, Virtual Brain Cloud No. 826421, Ministero Sviluppo Economico; Contratto di sviluppo industriale "Farmaceutica e Diagnostica" (CDS 000606) and "NextGenerationEU" (project IR0000011, EBRAINS-Italy).

Declaration of Competing Interest

The authors declare that they have no competing financial interest that could have influenced the present paper.

Credit authorship contribution statement

Pierpaolo Sorrentino: Conceptualization, Writing – original draft, Writing – review & editing, Formal analysis, Supervision. **Emahnuel Troisi Lopez:** Conceptualization, Investigation, Writing – original draft, Writing – review & editing, Formal analysis. **Antonella Romano:** Writing – original draft, Formal analysis, Writing – review & editing. **Carmine Granata:** Software, Validation, Resources. **Marie Constance Corsi:** Writing – original draft, Formal analysis, Writing – review & editing. **Giuseppe Sorrentino:** Writing – review & editing. **Viktor Jirsa:** Supervision, Writing – review & editing.

Data availability

The availability of the data was not previously included in the ethical approval, hence it is not possible to share data without specific approval from the local ethic committee.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2023.120260.

References

Amico, E., Goñi, J., 2018. The quest for identifiability in human functional connectomes. Scientific reports 8, 1–14.

Andersen, L.M., Jerbi, K., Dalal, S.S., 2020. Can EEG and MEG detect signals from the human cerebellum? Neuroimage 215, 116817. doi:10.1016/j.neuroimage.2020.116817.

Beggs, J.M., Plenz, D., 2003. Neuronal Avalanches in Neocortical Circuits. J. Neurosci. 23, 11167–11177. doi:10.1523/JNEUROSCI.23-35-11167.2003.

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: A practical and powerful approach to multiple testing. Journal of the Royal Statistical Society: Series B (Methodological) 57 (1), 289–300.

Bullmore, E., Sporns, O., 2009. Complex brain networks: graph theoretical analysis of structural and functional systems. Nature reviews neuroscience 10, 186–198.

Cellier, D., Riddle, J., Petersen, I., Hwang, K., 2021. The development of theta and alpha neural oscillations from ages 3 to 24 years. Dev. Cogn. Neurosci., 100969 doi:10.1016/j.brainres.2018.04.007.

Chialvo, D.R., 2010. Emergent complex neural dynamics. Nature Phys 6, 744–750. doi:10.1038/nphys1803.

da Silva Castanheira, J., Orozco Perez, H.D., Misic, B., Baillet, S., 2021. Brief segments of neurophysiological activity enable individual differentiation. Nat Commun 12, 5713. doi:10.1038/s41467-021-25895-8.

Deco, G., Jirsa, V.K., McIntosh, A.R., 2011. Emerging concepts for the dynamical organization of resting-state activity in the brain. Nat Rev Neurosci 12, 43–56. doi:10.1038/nrn2961.

Demuru, M., Fraschini, M., 2020. EEG fingerprinting: Subject-specific signature based on the aperiodic component of power spectrum. Comput Biol Med 120, 103748.

Demuru, M., Gouw, A.A., Hillebrand, A., Stam, C.J., van Dijk, B.W., Scheltens, P., Tijms, B.M., Konijnenberg, E., ten Kate, M., den Braber, A., Smit, D.J.A., Boomsma, D.I., Visser, P.J., 2017. Functional and effective whole brain connectivity using magnetoencephalography to identify monozygotic twin pairs. Sci Rep 7, 9685. doi:10.1038/s41598-017-10235-y.

Donoghue, T., Haller, M., Peterson, E.J., Varma, P., Sebastian, P., Gao, R., Voytek, B., 2020. Parameterizing neural power spectra into periodic and aperiodic components. Nat. Neurosci. 23 (12), 1655–1665. doi:10.1038/s41593-020-00744-x.

Eickhoff, S.B., Yeo, B.T., Genon, S., 2018. Imaging-based parcellations of the human brain. Nature Reviews Neuroscience 19 (11), 672–686.

Finn, E.S., Rosenberg, M.D., 2021. Beyond fingerprinting: Choosing predictive connectomes over reliable connectomes. NeuroImage 239, 118254. doi:10.1016/j.neuroimage.2021.118254.

Fries, P., 2015. Rhythms for Cognition: Communication through Coherence. Neuron 88, 220–235. doi:10.1016/j.neuron.2015.09.034.

Friston, K.J., 2011. Functional and Effective Connectivity: A Review. Brain Connectivity 1, 13–36. doi:10.1089/brain.2011.0008.

Glerean, E. (2014). Mantel test - Matlab implementation. Retrieved from doi:10.6084/m9.figshare.1008724.v3.

Hillebrand, A., Tewarie, P., van Dellen, E., Yu, M., Carbo, E.W.S., Douw, L., Gouw, A.A., van Straaten, E.C.W., Stam, C.J., 2016. Direction of information flow in large-scale resting-state networks is frequency-dependent. Proceedings of the National Academy of Sciences 113, 3867–3872. doi:10.1073/pnas.1515657113.

Liparoti, M., Troisi Lopez, E., Sarno, L., Rucco, R., Minino, R., Pesoli, M., Perruolo, G., Formisano, P., Lucidi, F., Sorrentino, G., 2021. Functional brain network topology across the menstrual cycle is estradiol dependent and correlates with individual well-being. Journal of neuroscience research 99, 2271–2286.

Novelli, L., Razi, A., 2022. A mathematical perspective on edge-centric brain functional connectivity. Nature Communication 13, 2693. doi:10.1038/s41467-022-29775-7.

Oostenveld, R., Fries, P., Maris, E., Schoffelen, J.-M., 2010. FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. Computational Intelligence and Neuroscience, e156869 doi:10.1155/2011/156869, 2011.

Pesoli, M., Rucco, R., Liparoti, M., Lardone, A., D'aurizio, G., Minino, R., Troisi Lopez, E., Paccone, A., Granata, C., Curcio, G., 2022. A night of sleep deprivation alters brain connectivity and affects specific executive functions. Neurological Sciences 43, 1025–1034.

Polverino, A., Lopez, E.T., Minino, R., Liparoti, M., Romano, A., Trojsi, F., Lucidi, F., Gollo, L., Jirsa, V., Sorrentino, G., Sorrentino, P., 2022. Flexibility of Fast Brain Dynamics and Disease Severity in Amyotrophic Lateral Sclerosis. Neurology doi:10.1212/WNL.0000000000201200.

Porta, L.D., Copelli, M., 2019. Modeling neuronal avalanches and long-range temporal correlations at the emergence of collective oscillations: Continuously varying exponents mimic M/EEG results. PLOS Computational Biology 15, e1006924. doi:10.1371/journal.pcbi.1006924.

Rabuffo, G., Fousek, J., Bernard, C., Jirsa, V., 2021. Neuronal Cascades Shape Whole-Brain Functional Dynamics at Rest. eNeuro 8, ENEURO.0283-21.2021. doi:10.1523/ENEURO.0283-21.2021.

Romano, A., Trosi Lopez, E., Liparoti, M., Polverino, A., Minino, R., Trojsi, F., Bonavita, S., Mandolesi, L., Granata, C., Amico, E., Sorrentino, G., Sorrentino, P., 2022. The progressive loss of brain network fingerprints in Amyotrophic Lateral Sclerosis predicts clinical impairment. NeuroImage: Clinical 35, 103095. doi:10.1016/j.nicl.2022.103095.

Rucco, R., Bernardo, P., Lardone, A., Baselice, F., Pesoli, M., Polverino, A., Bravaccio, C., Granata, C., Mandolesi, L., Sorrentino, G., Sorrentino, P., 2020. Neuronal Avalanches to Study the Coordination of Large-Scale Brain Activity: Application to Rett Syndrome. Frontiers in Psychology 11.

Shriki, O., Alstott, J., Carver, F., Holroyd, T., Henson, R.N.A., Smith, M.L., Coppola, R., Bullmore, E., Plenz, D., 2013b. Neuronal avalanches in the resting MEG of the human brain. Journal of Neuroscience 33, 7079–7090. doi:10.1523/JNEUROSCI.4286-12.2013.

Sorrentino, P., Rucco, R., Baselice, F., De Micco, R., Tessitore, A., Hillebrand, A., Mandolesi, L., Breakspear, M., Gollo, L.L., Sorrentino, G., 2021a. Flexible brain dynamics underpins complex behaviours as observed in Parkinson's disease. Scientific reports 11, 1–12.

Sorrentino, P., Rucco, R., Lardone, A., Liparoti, M., Troisi Lopez, E., Cavaliere, C., Soricelli, A., Jirsa, V., Sorrentino, G., Amico, E., 2021b. Clinical connectome fingerprints of cognitive decline. NeuroImage 238, 118253. doi:10.1016/j.neuroimage.2021.118253.

Sorrentino, P., Seguin, C., Rucco, R., Liparoti, M., Troisi Lopez, E., Bonavita, S., Quarantelli, M., Sorrentino, G., Jirsa, V., Zalesky, A., 2021c. The structural connectome constrains fast brain dynamics. eLife 10, e67400. doi:10.7554/eLife.67400.

Sorrentino, P., Rabuffo, G., Baselice, F., Troisi Lopez, E., Liparoti, M., Quarantelli, M., Sorrentino, G., Bernard, C., Jirsa, V., 2023. Dynamical interactions reconfigure the gradient of cortical timescales. Network Neuroscience 7 (1), 73–85. doi:10.1162/netn_a_00270.

Troisi Lopez, E., Minino, R., Liparoti, M., Polverino, A., Romano, A., De Micco, R., Lucidi, F., Tessitore, A., Amico, E., Sorrentino, G., Jirsa, V., Sorrentino, P., 2023. Fading of brain network fingerprint in Parkinson's disease predicts motor clinical impairment. Human Brain Mapping 44, 1239–1250. doi:10.1002/hbm.26156.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., … Joliot, M., 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage 15 (1), 273–289.

Van De Ville, D., Farouj, Y., Preti, M.G., Liégeois, R., Amico, E., 2021. When makes you unique: Temporality of the human brain fingerprint. Science Advances 7. doi:10.1126/sciadv.abj0751, eabj0751.

Zalesky, A., Fornito, A., Cocchi, L., Gollo, L.L., Breakspear, M., 2014. Time-resolved resting-state brain networks. Proceedings of the National Academy of Sciences 111, 10341–10346. doi:10.1073/pnas.1400181111.

Zamani Esfahlani, F., Jo, Y., Faskowitz, J., Byrge, L., Kennedy, D.P., Sporns, O., Betzel, R.F., 2020. High-amplitude cofluctuations in cortical activity drive functional connectivity. Proceedings of the National Academy of Sciences 117, 28393–28401. doi:10.1073/pnas.2005531117.