Successful response inhibition corresponds to a network reshuffling of hierarchically organized cortical modules

Giampiero Bardella[1,2,3], **Franco Giarrocco**[1,2], **Emiliano Brunamonti**[2], **Pierpaolo Pani**[2], **and Stefano Ferraina**[2,*]

[1]PhD Program in Behavioral Neuroscience, Sapienza University of Rome, Piazzale Aldo Moro 5, 00185, Rome, Italy
[2]Department of Physiology and Pharmacology, Sapienza University of Rome, Piazzale Aldo Moro 5, 00185, Rome, Italy
[3]INFN Sezione di Roma, Department of Physics, Sapienza University of Rome, Piazzale Aldo Moro 5, 00185, Rome, Italy
[*]corresponding author(s): Stefano Ferraina (stefano.ferraina@uniroma1.it)

ABSTRACT

Despite recent works have investigated functional and effective cortical networks in animal models, the dynamical information transfer among functional modules underneath cognitive control is still largely unknown. Here, we addressed the issue using Transfer Entropy and graph theory methods on neural activities recorded from a multielectrode array in the dorsal premotor cortex of rhesus monkeys. We focused our analysis on the decision time of a Stop-signal (countermanding) task. When comparing trials with successful inhibition to those with generated movement the local network resulted organized in four classes of modules hierarchically arranged and differently partaking in information transfer. Interestingly, the hierarchical organization of modules changed during the task, being different for generated movements and cancelled ones. Our results suggest that inhibitory control is based on a topological re-organization of the premotor functional network.

1 Introduction

The brain is a complex system formed by billions of neurons that, being densely interconnected give rise to networks organized in modules of different dimensions. At the small-scale modules correspond to single neurons, at the large scale they are great-specialized brain areas. In between, at the mesoscale level, modules are aggregates of neurons (populations) of different size (columns; specialized sub-regions; etc).

In the last fifty years, neuroscience has tried to describe brain computations by linking neural activities to behavior. Crucial in this process, at every scale of investigation, is the understanding of how different modules interact and how information is shared and processed among parts.

In this context, the neurophysiological approach to brain functions with recording microelectrodes, performed mainly in animal models[1] provided invaluable advances. Indeed, the high spatial resolution of the method proved to be suitable for linking neurons activity to behavior, to describe the organization of local microcircuits and, sometimes, of the over standing larger networks[2–4].

Most of these studies referred to the analysis of single unit activity, others focused more on mesoscopic signals as the local field potentials (LFPs), indicating the average synaptic input to the explored area (for a review see[5]) and, to a lesser extent, on signals sampling the average spiking activity of discrete population[6–8].

Recent advances on microelectrode fabrication allowed us to sample sufficiently large populations of

neurons, locally in the studied brain regions, and to approach the investigation of their collective dynamics. Indeed, it has been shown that neurons express more of their contribution to complex behavioral functions either when observed as coordinated functional ensembles[9–17] or described as common responses to the input they receive, supporting the modular organization proposed by pioneering studies[18].

In the motor domain, micro and mesoscopic (LFPs) descriptions have largely proved that the dorsal premotor (PMd) cortex of primates is linked to motor planning maturation and decision making[19–22].

However, except for the cosine population code introduced by Gergopoulos and colleagues about 40 years ago[23,24], most population-based descriptions have been proposed after the advent of the simultaneous multiple electrode recordings and mainly for single unit spiking activity.

Here, we studied the PMd's local dynamics using a mesoscopic measure of spiking activity (SA, see Section 4.4) and by implementing a combined information-theory and graph-theory approach to describe how the collective interaction of recorded modules is linked to a successful inhibitory control of motor responses, one of the core executive functions.

2 Results

We investigated, at the mesoscopic scale, the information transfer and functional directed connectivity patterns among discrete populations of neurons during the motor decision phases of arm movements. The task the animals were engaged in (Fig.1) required either to move the arm toward a peripheral target (**Go trials**) or to cancel the movement in case of a less frequent appearance of a Stop signal (**Stop trials**; see 4.2).

The two types of trials were randomly presented. During Go trials, after the disappearance of the central target (Go signal) the monkeys were instructed to reach the peripheral target to obtain the reward. In Stop trials, after the Go signal, the central target reappeared (Stop signal) after a variable delay, called SSD (Stop signal delay). In these trials the monkeys were required to refrain from moving to earn the reward (correct Stop trials). If the monkey were unable to stop, the trials were classified as wrong Stop trials, and no reward was provided. Since the SSDs were varied according to a staircase procedure based on the performance, correct Stop trials constituted approximately 50% of Stop trials (see Section 4.2 and Table 1). This task allows us to estimate the behavioral time window within which the decision to move (or to stop) is taken. This interval is commonly referred to as the Stop signal reaction time (SSRT). Table 1 shows the behavioral parameters of the 21 recording sessions analyzed in the present study (12 for monkey P; 9 for monkey C). The average reaction time (RT) was 656 ms (129 ms SD, SD standard deviation) for Monkey P and 632 ms (119 ms SD) for Monkey C and the average SSRT was 217 ms (34 ms SD) for Monkey P and and 213 ms (26 ms SD) for Monkey C. Fig.2 shows, for one example session (S6) of Monkey P, the time evolution of the SA obtained from each cortical module explored separately for Go trials (dark green traces), correct Stop trials (red traces) and wrong Stop trials (black races) in the analysis epoch T, including the SSRT (see Figure 2 caption for more details). For the session shown in Fig.2 (for all other sessions see Table S1) we found (multi comparisons with Bonferroni correction) that 78% of the cortical modules displayed a significant difference of activity between Go and correct Stop conditions after the Stop signal presentation (i.e., during the specific SSRT). In the same interval we found that 65% of the modules show a significant difference between correct and wrong Stop conditions. None of the modules showed significant difference between latency-matched Go and wrong Stop condition. For the different conditions of the performed task, there was a significant modulation of the SA sampled every 400 micrometers across all sessions, thus confirming previous observations on the role of PMd in inhibitory motor control. However, at this stage, nothing could be inferred about the information transfer and the functional relationship between different modules.

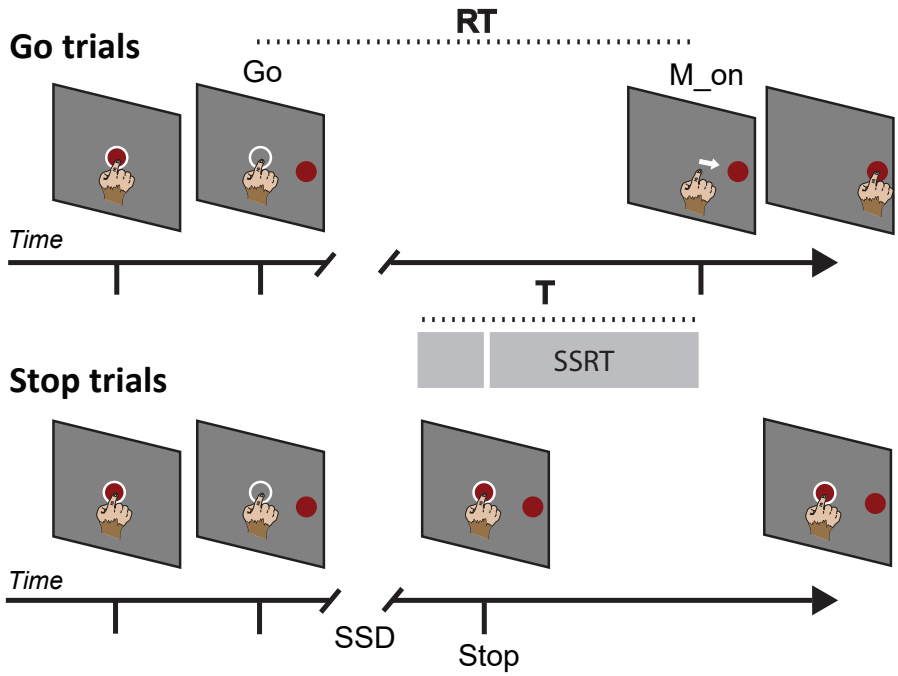

Figure 1. Sequence of behavioural events characterizing the task. Go and Stop trials were randomly intermixed during each session. The epoch (T) of neural analysis is shown as a grey bar. For Go trials, the SSRT marks the time, before movement execution, that would have corresponded to the presentation of the Stop signal. For correct Stop trials SSRT marks the time, after the Stop signal presentation, that would have corresponded to the movement execution. RT, reaction time; SSD, Stop signal delay; SSRT, Stop signal reaction time. White circles are feedbacks for the touch provided to the animals. Go: time of Go signal appearance; Stop: time of Stop signal appearance. M_on: time for Movement onset.

2.1 Local spiking activities can be grouped in different functional classes

To study the information transfer between recorded PMd modules in the epoch preceding the decision to perform/cancel the movement we used Transfer Entropy (TE), a well-established model-free information theoretic measure. To extract the TE values for each module, we adopted a fixed epoch duration (T= 400 ms) among animals and recording sessions which allowed to include the estimated SSRT duration for each recording session as well as the last portion of the SSD (see Fig.1, grey bar and Section 4.3 for further details). We computed TE between the single-trial spiking activity time series, in the epoch T, independently for Go, correct Stop, and wrong Stop trials. In a given epoch, TE can detect asymmetric information flow among modules allowing them to be classified as drivers (or sources) or targets of information transfer (see 4.5 for details). Fig.3 shows the TE matrix for each trial type obtained for the data in Fig.2. A TE value > 0 indicates a statistically significant driver$\rightarrow$ target relation for the pair of modules considered. As deducible from the dense horizontal stripes, that vary between Go and correct Stop trials, a few of the modules are able to influence the activity of many others depending on the behavioral condition. To exhaustively describe the differences in the information transmission at the network level (i.e., the network topology) we resorted to graph-based measures. In this framework, each module is interpreted as a node of the network and each TE value as a link (or connection, see Section 4.6 for further details). We computed the vertex degree (VD), i.e. the number of connections per module. A high VD score suggests that a module is functionally linked to many others. When VD is low the opposite

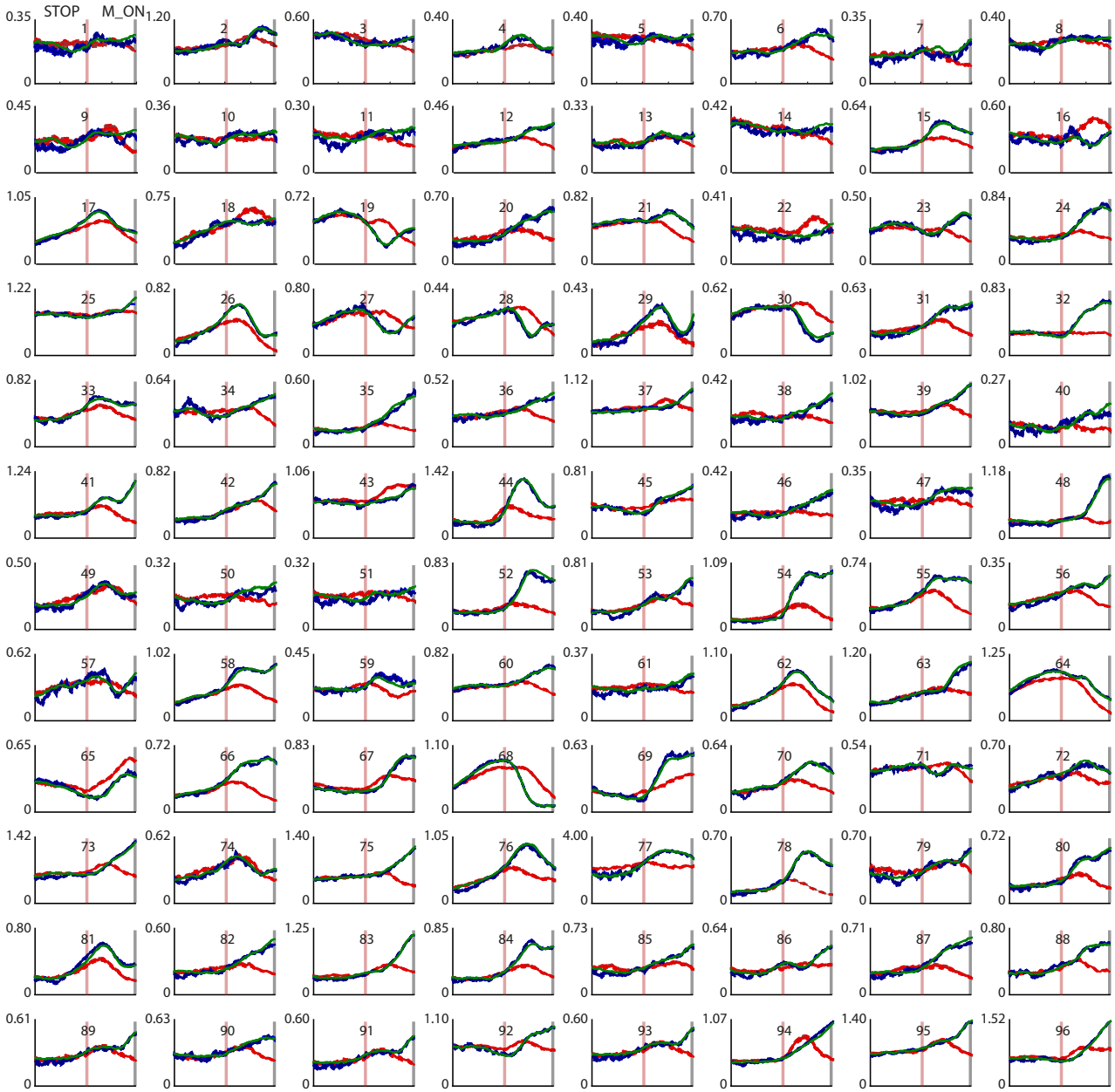

Figure 2. Neuronal SAs (mean ± SEM) in the three behavioural conditions for all modules of S6 of Monkey P. Dark green traces show the average activity during Go trials aligned to Movement onset (gray vertical line). Dark blue traces show the average activity during wrong Stop trials aligned to Movement onset (gray vertical line). Red traces show the average activity during correct Stop trials aligned to the Stop signal presentation (light red vertical line). The epoch between the Stop signal and the movement onset is the session SSRT. For each plot the Y axes are units of SAs while the X axes are coincident with the epoch T (see Figure 1); ticks in the first row left panel are every 100 ms.

Behavioural Results							
Monkey P							
S	$\overline{RT}_{Go}$	$\overline{RT}_{Wr}$	$\overline{SSD}_{CS}$	$\overline{SSD}_{Wr}$	$SSRT$	$P_{inhibit}$	*p-value*
1	590 ms	559 ms	273 ms	337 ms	285 ms	0.52	$p < 0.05$
2	584 ms	565 ms	305 ms	408 ms	227 ms	0.50	$p < 0.05$
3	618 ms	592 ms	335 ms	430 ms	238 ms	0.52	$p < 0.05$
4	573 ms	549 ms	301 ms	401 ms	221 ms	0.50	$p < 0.01$
5	572 ms	540 ms	293 ms	379 ms	236 ms	0.50	$p < 0.05$
6	643 ms	625 ms	402 ms	496 ms	195 ms	0.51	$p < 0.05$
7	600 ms	568 ms	340 ms	441 ms	209 ms	0.49	$p < 0.01$
8	664 ms	641 ms	469 ms	525 ms	170 ms	0.57	$p < 0.01$
9	786 ms	753 ms	528 ms	627 ms	213 ms	0.55	$p < 0.01$
10	672 ms	619 ms	421 ms	500 ms	216 ms	0.57	$p < 0.01$
11	766 ms	721 ms	497 ms	504 ms	230 ms	0.51	$p < 0.01$
12	868 ms	853 ms	675 ms	746 ms	160 ms	0.55	$p < 0.05$
Monkey C							
1	598 ms	523 ms	322 ms	383 ms	250 ms	0.57	$p < 0.01$
2	539 ms	460 ms	382 ms	378 ms	158 ms	0.60	$p < 0.05$
3	561 ms	522 ms	318 ms	403 ms	207 ms	0.58	$p < 0.05$
4	672 ms	631 ms	423 ms	507 ms	217 ms	0.60	$p < 0.05$
5	638 ms	612 ms	401 ms	512 ms	189 ms	0.56	$p < 0.05$
6	579 ms	533 ms	308 ms	426 ms	224 ms	0.60	$p < 0.01$
7	667 ms	620 ms	383 ms	498 ms	240 ms	0.59	$p < 0.05$
8	697 ms	672 ms	430 ms	546 ms	223 ms	0.60	$p < 0.05$
9	688 ms	662 ms	401 ms	547 ms	230 ms	0.60	$p < 0.05$

Table 1. Behavioural results. S, index of the recording session. $\overline{RT}_{Go}$, mean reaction time of Go trials. $\overline{RT}_{Wr}$, mean reaction time of wrong Stop trials. $\overline{SSD}_{CS}$, mean SSD of Correct Stop trials. $\overline{SSD}_{Wr}$, mean SSD of Wrong Stop trials $SSRT$, Stop signal reaction time. $P_{inhibit}$, inhibition probability. The *p-values* result from the independence (Kolmogorov-Smirnov test between RT_{Go} and RT_{Wr} distributions).

is true. Because of the asymmetry of TE, it is possible to discriminate between information directed to other modules (VD_{out}) and information received from other modules (VD_{in}). We examined the VD_{out} and the VD_{in} distributions for each recording session of both monkeys and we observed, for all sessions and conditions, that only the VDout distributions were fat-tailed (high variance; see supplementary Figure S1). The high values of VD_{out} associated with the tails suggest the presence of modules with many connections compared to the average. These modules are commonly defined as network hubs[25,26] (see Section 4.6). Thus, the analysis of the VD distributions suggested that the network had a more centralized organization when considering output connections than input ones. We labeled the hubs according to the behavioral condition in which they emerged: some hubs were present in all trials (**Super_drivers**); others in the trials when the movement was generated (**Go_drivers**) or successfully canceled (**Stop_drivers**) only. All the remaining modules, never acting as hubs, were classified as **Neutrals**. All sessions considered, the percentage of Hubs was about one quarter of the total modules. Super_drivers and Go_drivers were

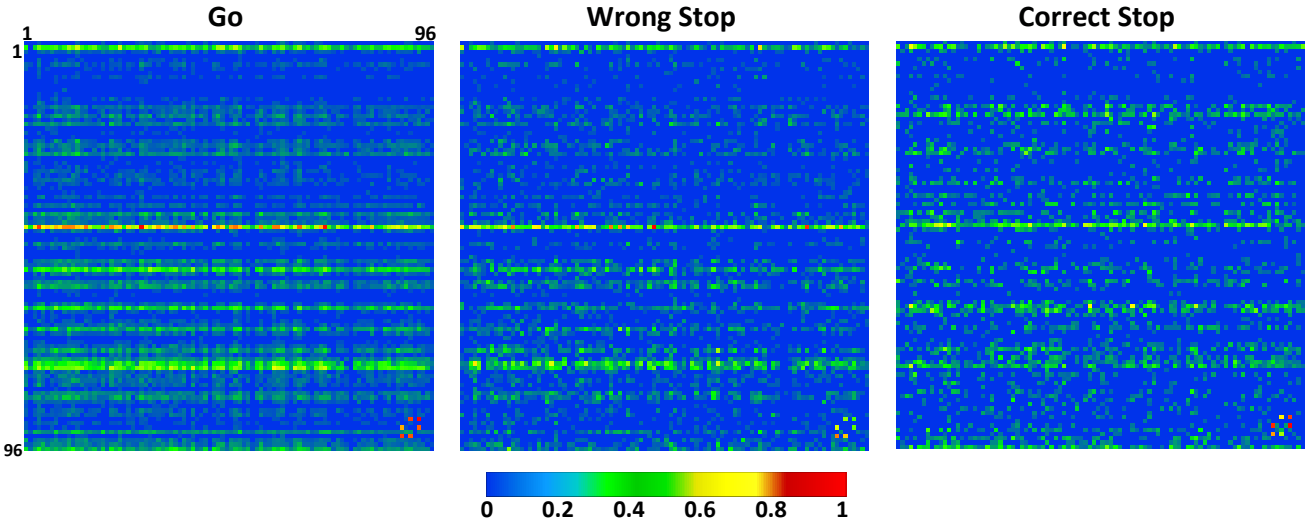

Figure 3. Average TE matrices of Go, wrong and correct Stop trials for data in Fig.2. Each entry of the matrix is color-coded according to the corresponding average TE value across trials. Statistically significant driver→ target relationships correspond to TE values > 0 for the pair of modules considered. Here, for graphical purposes only, TE is normalized to the maximum across behavioural conditions so to obtain TE $\in [0, 1]$. The scale bar indicates the normalized values of TE.

similarly expressed. The average percentage of Stop_drivers was the smallest in both animals (see Table 2 for the session average distribution of classes).

The presence of these four different classes of modules (3 different drivers and the Neutrals) straightforwardly shows that the intrinsic composition of the PMd information network is heterogeneous, with some of the modules operating as driving hubs only in relation to a specific behavioral outcome (moving vs withholding). This hints that the network configuration underlying information transmission is fine-grained and changes according to the decision leading to a specific behavioral output.

Fig.4 shows the topology of the PMd network in the three behavioral conditions for data in Fig.3 (see also supplementary Figure S2). Each connection indicates the presence of a significant TE value (i.e. a non-zero entry in the TE matrices); the number of outgoing connections of each module reflect its VD_{out} value, i.e., the number of modules on which it acts as a driver. Modules are labeled (see Figure 4 caption) according to the belonging class. Here, to highlight the relevant features of the classes changing between conditions only the links for the Go_drivers and the Stop_drivers are reported (for the complete networks including also the Super_drivers see Figure S2). The centralized organization is clearly visible in Go trials (and wrong Stop trials), confirming prior observations about the need for this state to facilitate movement generation[16]. Moreover, it is clear how the hubness of the network changes when movements are successfully canceled. Therefore, the emerging scenario is of a network preserving its constituent features (the presence of hubs) but changing in relation to the needed behavioral control. To this respect, Fig. 4 (and supplementary Figure S2) show how the Go_drivers (and Super_drivers) are hubs during Go and wrong Stop trials while the Stop_drivers emerge as important nodes of information spreading in correct Stop trials only. For other example sessions see Figure S3.

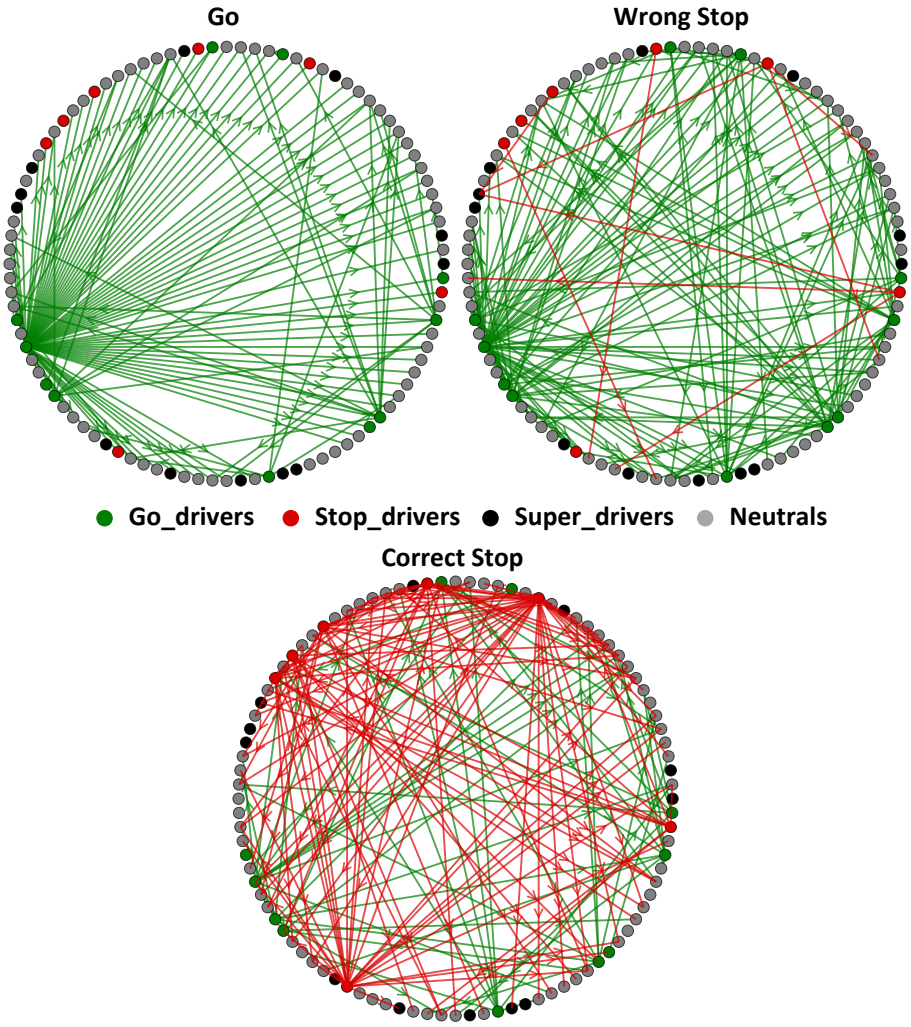

Figure 4. Information network of Go, wrong and correct Stop trials for data in Fig.2 and Fig.3. Each node and the respective connections are colour coded accordingly to the corresponding class (see legend in the figure). Nodes with an high number of outgoing connections are the information-spreaders hubs (see text for details). Here, for illustrative purposes and to better highlight the differences between behavioural conditions, the connections of the Super_drivers are removed and only the 10% of strongest connections are shown (for the complete version including the Super_drivers see supplementary Figure S2). The coordinates of each node on the circle are preserved for each network. For other examples of single session see figure S3.

2.2 The hierarchical organization of PMd modules changes when movements are successfully canceled

The VD_{out} values and the direction of information flow identified by TE allow establishing a hierarchy of information transmission among modules. Indeed, according to the TE-based definition of drivers and targets, the past of a driver gives statistical significant contribution to the prediction of the present state of a target. This puts the drivers hierarchically above the targets. To this extent, the existence of network hubs (i.e., the classes) means that a few modules determine the state of many others depending on the behavioral condition. As a result, depending on the behavioral condition, the classes selectively regulate the network's overall configuration.

Classes Composition	
Monkey P	$N = 96$
Class	$\boldsymbol{\mu \pm SEM}$
Go_drivers	9.3 ± 1.06
Stop_drivers	3.6 ± 1.4
Super_drivers	11.8 ± 1.6
Neutrals	71.3 ± 2.3
Monkey C	$N = 96$
Class	$\boldsymbol{\mu \pm SEM}$
Go_drivers	7.4 ± 0.8
Stop_drivers	4.1 ± 0.6
Super_drivers	9.4 ± 1.1
Neutrals	75.0 ± 1.8

Table 2. Classes composition. For each monkey the composition of classes averaged over recording sessions is reported. Composition is expressed as the average number of nodes (μ) belonging to each class. SEM, standard error of the mean. N, the number of modules available.

Fig.5 shows that, considering all sessions (see supplementary figure S4 for the single sessions), the Super_drivers had the highest values of VD_{out} compared to other classes in all three behavioral conditions, thus resulting as the primary information-spreaders hubs. Therefore, the Super_drivers are located at the highest hierarchical level in the network as they regulate information transfer whatever decision, moving or stopping, is taken (see also next paragraph).

Conversely, the Go_drivers and the Stop_drivers displayed a different engagement (lower VD_{out} values) in Go and correct Stop trials, suggesting that their hierarchical role changes in relation to the type of motor decision being processed.

Indeed, the Stop_drivers were below the Go_drivers in the Go trials and the Go_drivers were below the Stop_drivers during correct Stop trials. Interestingly, the Go_drivers had higher degree during wrong Stop trials compared to correct Stop trials while the opposite holds for the Stop_drivers. This confirms, at the level of information processed within the network, what was stated above: during wrong Stop trials a neural processing halfway between Go and correct Stop trials but more akin to that of Go trials unfolds.

To further clarify the hierarchical dependencies between classes, we examined the VD_{out} as a function of TE. This gives a measure of the classes' robustness (see Section 4.6). Indeed, for a class that is functionally relevant to a specific behavioral condition we expect not only a high number of connections (high VD_{out} value) but also that these connections have a high information content (high TE value). Results for the robustness analysis are reported in Fig.6. As expected, the Super_drivers proved to be the class of highest hierarchical level in all three behavioral conditions. In fact it showed the highest value of VD_{out} as a function of the information processed, i.e. the highest robustness. Go_drivers and Stop_drivers then followed in this scale, as they played a complementary role in the Go and in the wrong Stop trials and in the correct Stop trials respectively

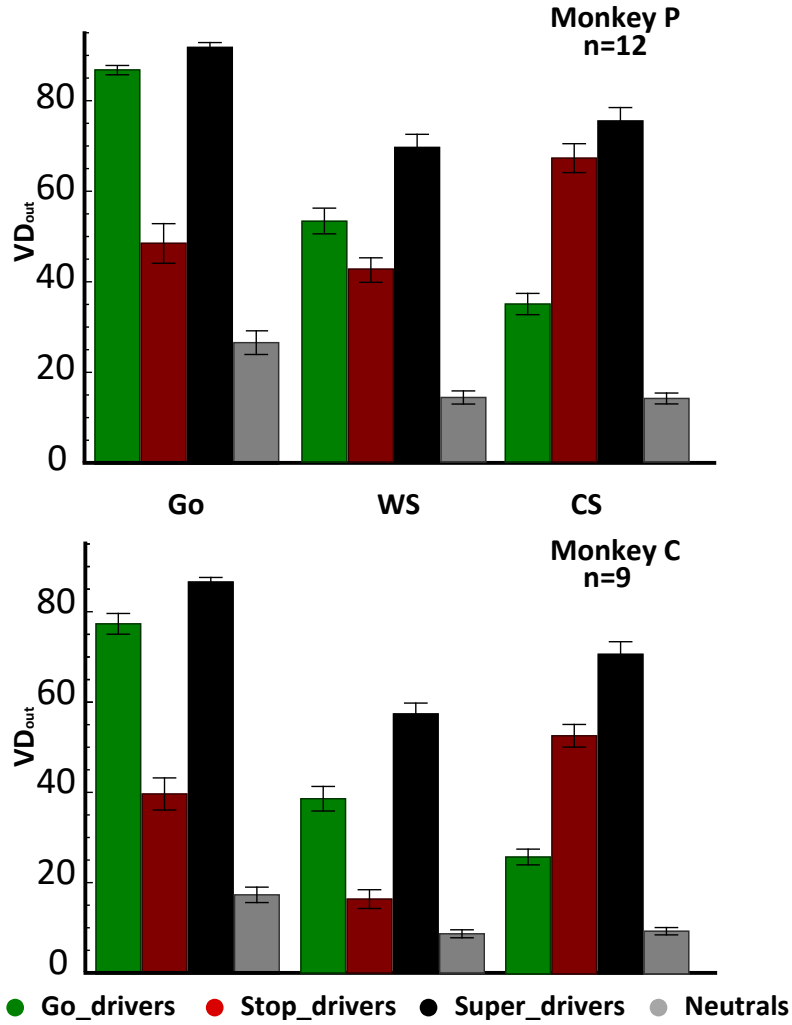

Figure 5. Topology of information transmission. VD_{out} values averaged over recording sessions compared between classes and across behavioural conditions (**Go**: Go trials; **WS**: wrong Stop trials; **CS**: correct Stop trials; **n**: number of recording sessions). The Super_drivers show the highest values of VD_{out} compared to other classes in all behavioural conditions. Go_drivers and Stop_drivers show the second highest VD_{out} values during both Go and wrong Stop trials and correct Stop trials respectively. Colours reflect the neuronal classes as in Figure 4. Error bar is the SEM over recording sessions. Statistics is based on one-way ANOVA within behavioural conditions with all $p \leq 0.05$ Bonferroni corrected for multiple comparison.

2.3 Different interactions among neuronal classes characterize behavioural conditions

To summarize the interactions among classes we computed the average amount of information exchanged between the four classes during behavioral conditions. To this end, we constructed an interaction **index I** proportional to the total intensity of the information exchanged between the classes. We condensed I into a 4x4 matrix that we then represented as a network in which each node is a class and connections are the weighted sum of TE values (see 4.6). This makes it possible to have a compact picture of the differences between behavioral conditions in terms of the interactions between classes.

Fig.7 shows the results averaged over recording sessions (see Table S2). The Super_drivers were confirmed to be the highest order class in the network since they send out the largest amount of information

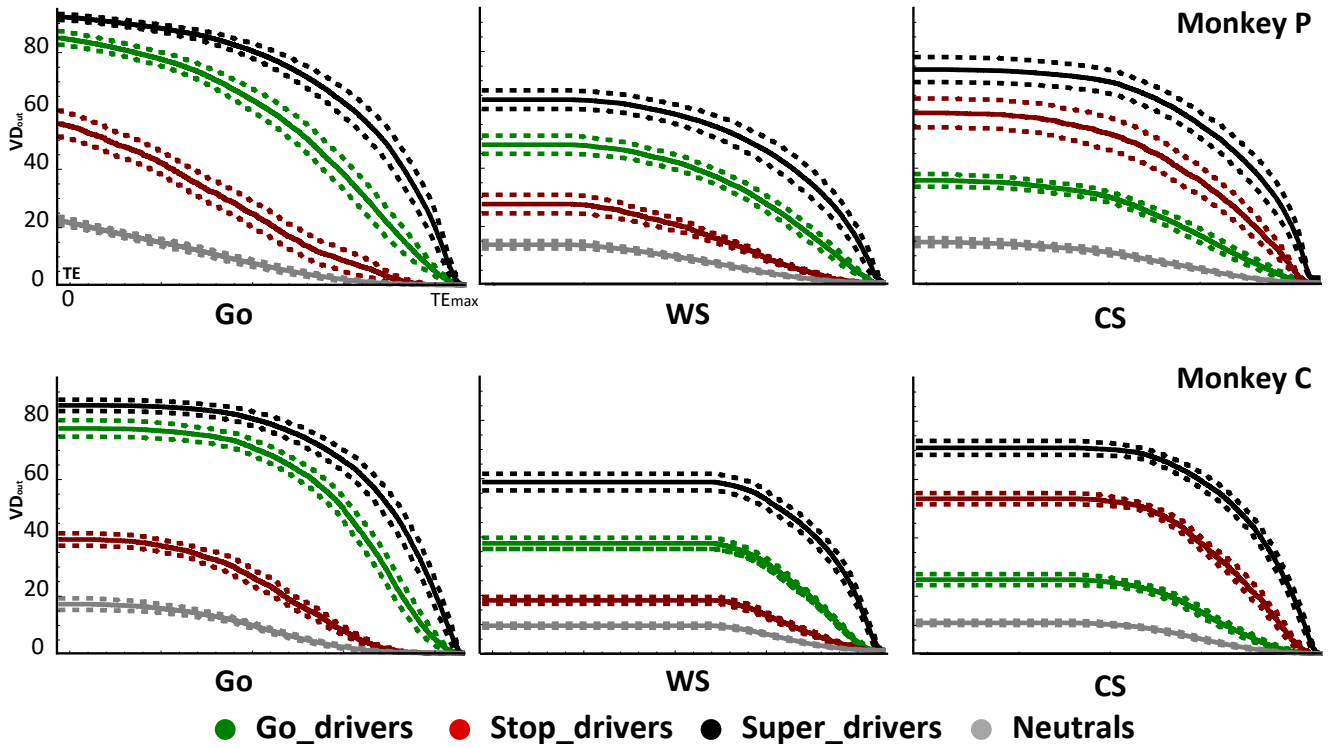

Figure 6. Classes robustness. Robustness curves averaged (mean ± SEM) over recording sessions compared between classes and across behavioural conditions (**Go**: Go trials; **WS**: wrong Stop trials; **CS**: correct Stop trials;)

in all behavioral conditions compared to other classes. This indicates that the extent of communications of the Super_drivers with the network was significantly greater than that of the Go_drivers and the Stop_drivers. This implies that the Super_drivers determine the global state of the network with the Go_drivers and the Stop_drivers playing a supporting role depending on the behavioral outcome. The latter indeed distributed the highest information content in Go and correct Stop trials respectively. It is interesting to highlight how in Go trials the highest value of I, besides the one from the Super_drivers, was the one going from the Go_drivers to the Neutrals. In comparison, the contributions from the Go_drivers to the Stop_drivers and to the Super_drivers were negligible. The exact complementary situation was found in correct Stop trials, with the highest value of I, besides the one from the Super_drivers, detectable from the Stop_drivers to the Neutrals. It is worth noting that the direction of the interactions during behavioral conditions corroborated the specialization of the Go and Stop drivers. Indeed, during the Go trials information traveled from the Go_drivers to the Stop_drivers while the opposite happened during correct Stop trials. Combining these results with those obtained from the previous analyses, it is clear how Go_drivers and Stop_drivers are hierarchically at a lower level than the Super_drivers: not only do they express fewer connections but also the amount of influence they exert on other classes (the weight of the connections, i.e. TE) is less than that of the Super_drivers (see also Fig.5 and Fig.6).

The configuration observed during wrong Stop trials was, as expected, a hybrid between Go and correct Stop trials with more features similar to the Go than that to the correct Stop. Indeed, like in the Go trials the highest contributions to the information processed came from the Super_drivers and the Go_drivers. However, some features of correct Stop trials, such as the contributions from the Stop to the Go and Super drivers, were also present. To sum up, we demonstrated that the TE-defined classes are hierarchically

organized in PMd during movement planning and suppression and that information is processed differently during correct Stop trials compared to Go trials. Results revealed the Super_drivers as the most topological central nodes in the network with the Go_drivers and the Stop_drivers playing a crucial supporting role in the processing of information during movement planning and inhibition respectively.

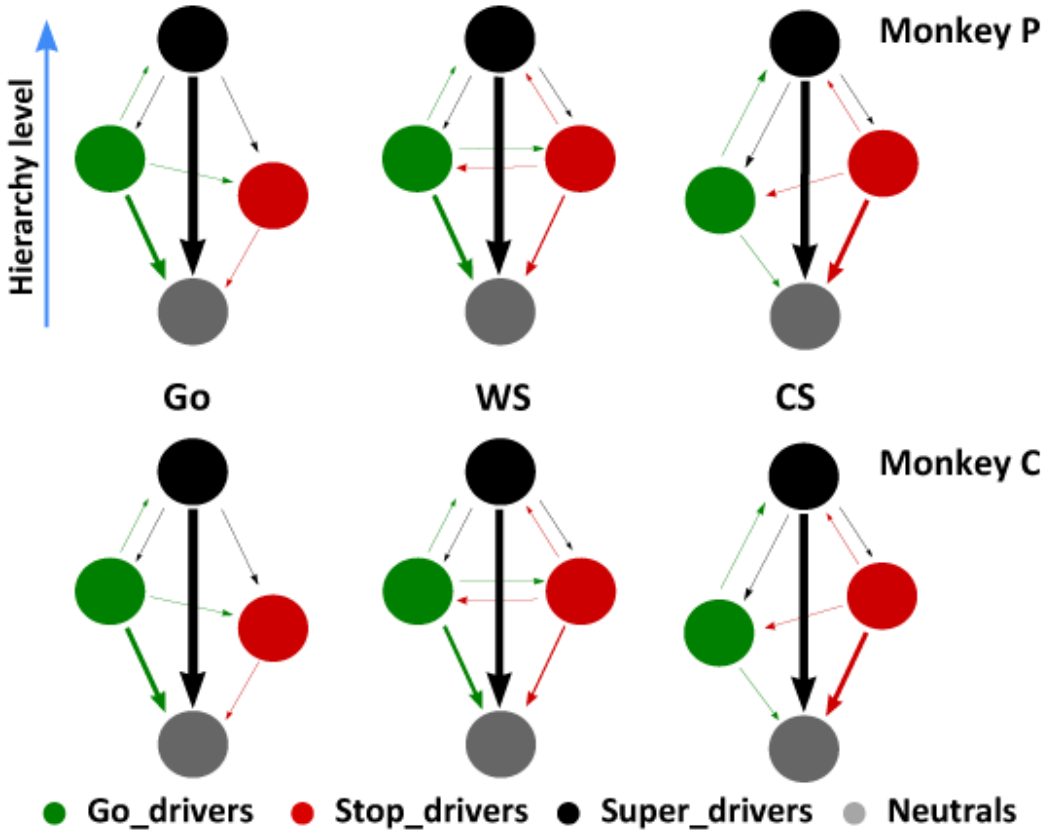

Figure 7. Network representation of interactions between classes: Colours codes for the classes are the same of the previous figures. The thickness of the arrows is scaled according to the corresponding value of **I**. The strength of I determines the level of hierarchy characteristic of every class, which is graphically represented as the vertical positioning of the nodes.

3 Discussion

In this work we investigated the patterns of local information transfer in a cortical network (in the PMd) directly involved in movement decision-making. We used a combined Transfer Entropy and graph theory-based approach to analyze simultaneously recorded spiking activities (from up to 96 channels). Our results contribute to moving forward the knowledge on the neural basis of motor control at different levels.

3.1 A topological approach to the organization and spreading of local information in a motor decision task

The first level of advancement is methodological: we employed a graph theory approach combined with information theoretic measures (specifically Transfer Entropy) to investigate neuronal interactions underlying motor control. Although TE is growing in popularity in modern neuroscience its application to invasive electrophysiological data has been so far very limited and restricted mostly to either single

neurons or in vitro[27–29] and in silico studies[30]. Shimono and Beggs[27] used it at the single neuron level to investigate the structure at different scale of rodent somatosensory cortex[27,31]. Timme and colleagues[32] recorded the activity of hundreds of neurons in cortico-hippocampal slice cultures and used TE to study the information transfer changes therein. An early contributions to this research topic is from Gerhard et al[33] that inspected the topology of spike trains recordings from the visual system of a rhesus monkey during a fixation task. However, the authors used a different approach to measure directed connectivity (the use of TE on invasive electrophysiological data from the visual system is also attributable to Besserve and collegues[34] which however did not resort to graph theory techniques). Another contribution comes from the work of Honey[35], that investigated a large-scale interregional anatomical network of the macaque cortex trough Transfer Entropy. An attempt to study voluntary action control through analysis of directed connectivity was made by Jahfari and colleagues[36] but on human MRI data. Hence, to the best of our knowledge, this report is one of the very few studies that uses graph theory to analyse the information transfer network of a specific cortical area at the mesoscale level in vivo and during a behavioural task. The level of description here obtained is more detailed compared to previous works. Indeed, we were able to specify how the decision on whether to move or to stop is implemented in PMd at the population level and who are the (key) players that manage information transmission. Notably, in our framework neither any a priori assumption nor a specific neural modelling technique was needed. Our completely data-driven approach, in addition to complement the most recent models for motor generation and suppression[37,38], permits to overcome their main limitation which resides in the requirement of many biophysical parameters to be tweaked and tuned before fitting with acceptable accuracy the experimental data. Although is still not largely used in behavioural neurophysiological studies at the small and mesoscale, a graph theory-based conceptualization of neural interactions, united with information theoretic measures, can be very profitable also compared to other common approaches relying on the analysis of covariance between neurons or mesoscopic signals[8,39–42] and should be exploited more. In fact, these methods are not straightforward in distinguishing the specific contributions of single neurons (or discrete populations of neurons) to the topology of network dynamics, which is indeed the strength of our strategy. On one hand, this range of methods allows for a fine temporal description of neural variability but on the other, due to their nature, are neither capable to describe the information flow between neuronal actors nor to provide quantitative insights on the topology of network connections and their hierarchical organization. Without this spectrum of details the computational strategy underlying motor control (and neural circuitry computation in general) would be yet elusive. Recently, some authors have started to follow the joint information theory-complex networks approach but for now, to the best of our knowledge, only on cortico-hippocampal[32], somatosensory cortex slice cultures[27] and during anesthesia[43]. It is known that to fully understand the neural mechanisms behind motor control future research should focus on cortico-cortical and cortico-subcortical interactions through massive simultaneous recordings. In this scenario, a topological information-based approach would be unquestionably necessary to gain an overall view and elicit detailed insights.

3.2 More than two horses are racing during response inhibition control

The second level of advancement concerns the novelty of our results compared to other studies, especially those that focused on the possible interaction among different classes of neurons during motor decision. We found, in the characterized PMd network, that neuronal activities could be organized around four different classes and that they actively participate, even with different roles (i.e., through different topological attributes), both in movements execution and cancellation. This constitutes a step forward in the conceptualization of the neural processes at the base of movement generation since all the widely accepted models for inhibitory control of movements[37,38,44–47] are deduced from the analysis of data from

single electrodes recordings and are based on the interaction of only two modules (or class of neurons) often reported as Go and Stop units. We instead demonstrated that information is hierarchically transferred between more than two actors, with the Super_drivers acting as network hubs. This reflects the existence of an intrinsic high-order complexity in functional communications and organization at the population level, even in small portions of the cortex, during behavioural control regardless of which the nature of neurons in each class might be (e.g., excitatory or inhibitory neurons). Indeed, based only on the information exchanged by each local module, we managed to isolate both condition-specific and non specific neuronal classes. It is worth stressing that we drew our conclusion on the heterogeneity of neuronal classes in a completely data-driven and model-free fashion, and this strengthens the results. Additionally, we showed the details of how this transfer occurs at the population level and to what extent it is linked to different behavioural outcomes. Our picture integrates the current view because besides specific classes involved in the generation (Go_drivers) and inhibition (Stop_drivers) of movements, it establishes the existence of a high order class (Super_drivers) not proposed in other works. This highlights, for the first time at the mesoscale resolution, the existence of a fine-grained organization of neural assemblies at the population level that handle intra-area information flow. It is worth pointing out that usual methods of studying neural activity profiles are not sufficient to infer all aspects of such architecture. The Super_drivers are higher in hierarchy with respect to the others for two reasons. The first is because they transmit information to the whole network without receiving from inside the same network. Indeed, from the information theoretic point of view, this means that the state of the other classes can be better predicted by the state of the Super_drivers compared to that of the other classes. Thus, the state of the whole local network depends on the state of the Super_drivers. The second one is topological, being the Super_drivers the most widespread hubs across behavioural conditions. The found subdivision in classes, the presence of hubs and topological central nodes deputed to the rerouting of communications reveal that the cortical information dynamics behind motor control is extremely rich and cannot be entirely explained by the current proposed models. The found topology also implies that the presence of high-degree nodes is a constituent feature of neural processing in a cortical network directly involved in cognitive control, as is the PMd. This is consistent with our previous study[16], in which we showed how the functional PMd network organization differs between movement generation and inhibition in terms of hierarchy and centrality of nodes. It is also in agreement with other works that found fat-tailed degree distributions in silico[48], in cortical and hippocampal in vitro networks[27,49–52], in vivo[53] and anatomical structural networks[35]. We found that the arrangement of the PMd information network depends on the behavioural condition, passing from a centralized state during movement planning to a different one during movement inhibition, the latter characterized by high values of Betweenness Centrality and a minor transfer of information.

We interpret the reorganization as the execution in the local network of a command originating from other regions. Indeed, as known, the PMd is part of a larger network surbserving motor control based on frontal, parietal, subcortical and spinal structures. It is reasonable to think that during Go trials the hubs serve to convey the command to move to other (and possibly anatomically lower) cortical, subcortical, and spinal circuits that will eventually promote muscle activation. In this picture, the state observed during correct Stop trials could reflect the PMd collective reaction to the incoming inhibitory thalamic input that prevents the execution of the programmed movement. Therefore, the volition to inhibit would be locally implemented as 'the attenuation of the movement state', which seems convenient ad easy to implement at the network level by detouring information flow through an higher number of paths between nodes and decreasing the amount of information involved. Future studies will be necessary to investigate to whom the hubs project to. One weakness of this study is that we cannot account for the information dynamics between PMd and other structures of the reaching network. Therefore, additional research will be needed to unambiguously clarify these interactions. Lo et al.[38] also introduced a certain degree of hierarchical

organization in the form of a top-down control regulating the activation of the Go and Stop units. However, as also stated in Schall et al.[46], the control unit embodied in their model resembled an external homunculs endowed with the ability to tune the parameters to appropriately obtain the desired results. This marks a considerable difference with our report, in which, since our approach is completely data-driven, we did not need to adjust any external modelling unit to obtain the results. Conversely, we used it conceptually to contextualize our results in a wider circuitry frame. Lastly, our findings clearly show that hierarchical control is not only external but is also implemented locally by a specific neuronal class (the Super_drivers) over the others. Throughout the years, much evidence has been brought to support the idea that the brain is hierarchically organized both globally and locally on a spatial[16,54–65] (for a detailed review see Hilgetag et al., 2020[66]) and temporal scale[67–72]. As far as we know, this is the first work that deeply investigates the local hierarchy of a single cortical area known to have a crucial role in the motor system. These conclusions suggest that the collective network organization found in this work represents the neural implementation for the voluntary motor control at the PMd level.

4 Materials and methods

4.1 Subjects

Two male rhesus macaque monkeys (Macaca mulatta, Monkeys P and C), weighing 9 and 9.5 kg, respectively, were used. Animal care, housing, surgical procedures and experiments conformed to European (Directive 86/609/ECC and 2010/63/UE) and Italian (D.L. 116/92 and D.L. 26/2014) laws and were approved by the Italian Ministry of Health. Monkeys were pair-housed with cage enrichment. They were fed daily with standard primate chow that was supplemented with nuts and fresh fruits if necessary. During recording days, the monkeys received their daily water supply during the experiments.

4.2 Apparatus and task

The monkeys were seated in front of a black isoluminant background ($< 0.1cd/m2$) of a 17-inch touchscreen monitor (LCD, 800 x 600 resolution), inside a darkened, acoustic-insulated room. A non-commercial software package, CORTEX (http://www.nimh.gov.it), was used to control the presentation of the stimuli and the behavioural responses. Fig.1 shows the scheme of the general task: a stop-signal reaching task[73]. Each trial started with the appearance of a central target (CT) (red circle, diameter 1.9 cm). The monkeys had to reach and hold the CT. After a variable holding time (400–900 ms, 100 ms increments) a peripheral target (PT) (red circle, diameter 1.9 cm) appeared randomly in one of two possible locations (right/left) and the CT disappeared (Go signal). In Go trials, after the Go signal the subjects had to reach and hold the PT for a variable time (400-800 ms, 100 ms increments) to receive juice. Reaction times (RTs) were defined as the time between the presentation of the Go signal and the onset of the hand movement. In Stop trials, the sequence of events was the same until the Go signal. Then, after a variable delay (Stop signal delay SSD), the CT reappeared (Stop signal) and the monkeys had to hold the CT until the end of the trial, (800–1000 ms) to receive the reward (correct Stop trial). Conversely, removing the hand after the Stop signal constituted a wrong response (wrong Stop trial). The same amount of juice was delivered for correct Stop and correct Go trials. The inter-trial interval was set to 800 ms. Stop trials represented on average the 25% of all trials for Monkey P and 32% for Monkey C. To establish the duration of the SSDs, a staircase tracking procedure was employed. If the monkey succeeded in withholding the response, the SSD increased by one step (100 ms) in the subsequent Stop signal trial. Conversely, if the subject failed, the SSD decreased by one step.

4.3 Behavioural considerations

The stop-signal reaching task makes possible to calculate a behavioural measure that it is broadly considered as an index of efficiency in movement suppression: the Stop signal reaction time or SSRT. To estimate the SSRT the race model[74] is the accepted paradigm. This model describes the behaviour in the Stop trials as the result of two stochastic processes racing toward a threshold: the GO process triggered by the onset of the Go signal, which duration is represented by the RT, and the STOP process triggered by the onset of the Stop signal, which duration must be calculated. When the GO process wins the race the movement is generated (wrong Stop trials), alternatively it is withheld (correct Stop trials). The race model allows to estimate the SSRT by considering the duration of the GO process, the probability to respond and the SSDs. However, to make the race model applicable to study response inhibition, a central assumption must be satisfied: the GO process in the Stop trials must be the same as in the Go trials (independence assumption). Indeed, the RTs that are employed to estimate the SSRT are obtained from the Go trials distributions. To broadly validate this assumption, wrong Stop trials RTs must result shorter than the correct Go trials RTs[74] (see Table 1). To estimate the SSRT we employed the integration method because it has been proven to be the most reliable[75]. The method assumes that the finishing time of the Stop process corresponds to the nth Go RT, where n results from the multiplication of the ordered Go RTs distribution by the overall probability of responding $p(respond)$. The SSRT is then obtained by subtracting the average SSD from the nth Go RT. The SSRT can also be considered as the lead time that is required to inhibit a movement, or, simply, the time that precedes the start of a movement when a Stop signal, if presented, halts the generation of the same movement approximately 50% of the time. If the Stop signal is presented after this time, it will be less effective, because the neuronal phenomena that lead to movement generation will have already be started. If the Stop signal is presented before this time, it will be more effective in halting the movement. Consequently, the neuronal activity that is related to movement generation must occur before movement onset around the time that is defined by the SSRT. The aim of our study was to compare conditions in which a movement was planned and then generated (Go trials) to those in which a movement was planned and then inhibited (correct Stop trials). To correctly compare the two behavioural conditions, a time window (T) equivalent for both trial types must be defined. The duration of T corresponded to 400 ms. For Go trials, T was the last portion of the RT. For Stop trials, T allowed to include any SSRT duration computed in the different sessions (see Table 1) and a variable portion of the SSD (controlling for the duration of the shortest used). Behavioural parameters for the recording sessions of the two monkeys analyzed in this study are reported in Table 1.

4.4 Extraction and processing of neuronal data

A multielectrode array (Blackrock Microsystems, Salt Lake City) with 96 electrodes (spacing 0.4 mm) was surgically implanted in the left dorsal premotor cortex (PMd; arcuate sulcus and pre-central dimple used as references after opening of the dura) to acquire unfiltered electric field potentials (UFP; i.e., the raw signal) sampled at 24.4 kHz (Tucker Davis Technologies, Alachua, FL). As a measure of neuronal activity at the population level, local spiking activity (SA) was extracted offline from the raw signal, as in Mattia et al.[8], by computing the time-varying power spectra $P(\omega, t)$ from the short-time Fourier transform of the signal in 5-ms sliding windows. Relative spectra $R(\omega, t)$ were then obtained normalizing $P(\omega, t)$ by their average $P_{ref}(\omega)$ across a fixed window (30 minutes) for the entire recording. Thus, the average $R(\omega, t)$ across the $\omega/2\pi$ band $[0.2, 1.5]$ kHz represent the spectral estimated spiking activity. As better detailed in Mattia et al.[8], such estimate relies on two hypotheses. The first is that high ω components of the raw signal result from the convolution of firing rates $\nu(t)$ of neurons that are close to the electrode tip with a stereotypical single-unit waveform. The Fourier transform $K(\omega)$ of such an unknown waveform is canceled out in $R(\omega, t)$, which is therefore a good approximation of the ratio of firing rate spectra $|\nu(\omega,t)|^2$

$/|\nu(\omega,t)|_{ref}{}^2$. The second hypothesis is that high ω power $|\nu(\omega,t)|^2$ is proportional to the firing rate $\nu(t)$ itself[76], such that our estimated spiking activity is proportional to $\nu(t)$. As a last step, logarithmically scaled spiking activities were smoothed by a moving average (20 ms sliding window, 5ms step).

4.5 Quantifying information dynamics with Transfer Entropy

We first analysed the single-trials SA profiles of each module for each recording session for both animals. To remove noise and outliers from our data, we excluded from the analysis the trials for which the SAs showed peaks with an amplitude that exceeded the average of the activity by 2 standard deviations in the epoch of interest and for over 80% of the channels. This ensures that artifacts caused by non-physiological oscillations are excluded from the analysis. To examine the local information dynamics in the PMd, we constructed the single-trial information transfer networks for each behavioural condition of each recording session using multivariate Transfer Entropy (*TE*). The choice is due to the fact that TE is indicated (especially in its multivariate formulations) as more accurate compared to other metrics and is known to capture non-linear interaction in the system dynamic without assuming any particular model. Moreover, this measure is of growing interest in neuroscience and there is a thriving literature on it[?,77–80]. For its computation we adapted to our data-analysis pipeline some functions of the Matlab *MUTE* toolbox[81]. Given an ensemble of M time series, the multivariate information transfer from a *driver* time series X to a *target* time series Y, conditioned to the remaining $Z_{k=1,..,M-2}$ time series, can be quantified considering the present values of the target and the past values of both the driver and the Z through[77,81,82]:

$$TE_{X\to Y\,|\,Z} = H(Y_n \mid Y_n^-, Z_n^-) - H(Y_n \mid Y_n^-, X_n^-, Z_n^-)\ , \tag{1}$$

where Y_n is the vector that represent the present state n of Y, $X_n^- = [X_{n-1}, X_{n-2}, ...]$, $Y_n^- = [Y_{n-1}, Y_{n-2}, ...]$ and $Z_n^- = [Z_{n-1}, Z_{n-2}, ...]$ are the vectors that represent the past of X, Y and Z respectively. The vertical bar stands for conditional probability, e.g. $H(Y_n \mid Y_n^-, Z_n^-)$ is the entropy of the present state of Y conditioned to the knowledge of the past of Y and to the past of the remaining Z. H is the Shannon entropy[83], which in the case of Y is given by:

$$H(Y_n) = -\sum_n P(Y_n)\ log P(Y_n)\ , \tag{2}$$

where P indicates the probability density. Hence, using equation 2 expression 1 becomes

$$TE_{X\to Y\,|\,Z} = -\sum_n P(Y_n, Y_n^-, X_n^-, Z_n^-)\ log \frac{P(Y_n \mid Y_n^-, X_n^-, Z_n^-)}{P(Y_n \mid Y_n^-, Z_n^-)} \tag{3}$$

In this formulation *TE* grows if the past of the driver increases the information about the present of target more than the past of target itself and more than any other series contained in Z. Since the past of the driver is used to predict the present of the target, TE is not symmetric (i.e. $TE_{X\to Y} \neq TE_{Y\to X}$) and defines a direction in the information transfer. A crucial issue when estimating TE is the approximation of the vectors representing the past of the time series, a procedure known as *embedding*. The optimal embedding would be the one that includes only the components of X_n^-, Y_n^- and Z_n^- that are most informative in describing Y_n. Montalto and colleagues[81] described in details different procedures both for embedding and to evaluate the probability distribution functions needed to compute the entropy terms. We opted for

a non-uniform embedding scheme[84] paired with the computation of H based on kernels estimators[1]. In few words, the embedding method we chose iteratively selects components of the systems past based on a criterion for maximum relevance and minimum redundancy. In this context, maximum relevance means most significant in the sense of predictive information. Non-uniform embedding selects from the past of X, Y and Z only the components that are the most informative for the present of the target variable Y, progressively pruning non informative terms. The maximum number of past values, or maximum lag l, to consider for the pruning is fixed at the beginning of the procedure. The statistical significance is then progressively assessed cycling through the components of the past up to l and through the comparison with an null distribution built from the empirical values via a randomization procedure[81]. The components of the past of X,Y and Z are thus selected as statistical significant if they are significant above a desired level α. In our case the null distribution was obtained by 300 random shuffling of empirical values and we fixed $\alpha = 0.01$. Non-uniform embedding represents a convenient and more reliable[77] alternative to the common used approach known as uniform embedding that would select the past values X_n^-, Y_n^- and Z_n^- *a priori* and separately for each time series[81]. The probability density P needed for the computations of H was then calculated using kernel functions which weight the distance to the reference point to any other point in the time series and then average across all points. Such approach computes probabilities exploiting a local exploring of the state space and, importantly, has been proven to be more robust against unreliable estimations compared to other approaches[77]. Therefore, if at least one component from the past is selected by the non-uniform embedding procedure, the resulting $TE_{X \to Y \mid Z}$ is positive and statistically significant. When instead none of the components of the past provide statically significant information about the target the $TE_{X \to Y \mid Z}$ is exactly 0 and assumed non significant[81]. To avoid any further bias in the selection of the past values, we initially fixed l=50 ms, but, as expected, only a recent past was selected by the procedure, in line with similar studies[27,32]. Indeed, for each SA time series, a past no older than 10ms for each n of equation 2 was ever selected by the optimal embedding procedure.

4.6 Graph theoretical analysis

In our context the time series were the local SA recorded by the electrodes of the array. We computed $TE_{X \to Y \mid Z}$ (and $TE_{Y \to X \mid Z}$) with $Z_{k=1,..,M-2}$, for each pair of (X,Y) in the T epoch defined in Section 4.3 so to obtain a single-trial TE matrix for each behavioural condition (Go trials and correct and wrong Stop trials) for both monkeys. We interpreted the TE matrix as the adjacency matrix of a directed weighted network, in which the nodes are the single modules (electrodes) and the weighted edges are the $TE_{X \to Y \mid Z}$ (and $TE_{Y \to X \mid Z}$) with $Z_{k=1,..,M-2}$. To simplify the picture we ignored the main diagonal of the matrix thus excluding self-loops from the networks. Finding an optimal threshold to analyze network properties is a major challenge when dealing with connectivity matrices. To overcome the hindrance we used percolation as we did in our previous works[16,65]. Percolation is a tool that is rooted in statistical physics and since its first applications to the study of phase transitions of connected subgraphs in random networks[86], the method has progressively entered the toolbox of those neuroscientists interested in the hierarchical organization of brain networks[87–94]. During percolation the stability of network topology is monitored via the iterative removal of links based on the value of their weights. Therefore the process naturally discloses both network robustness properties and signatures of hierarchical self-organization simultaneously. Moreover, recent studies[88,92] have shown that percolation leads to identification of a specific network threshold that achieves the best balance between the removal of noise-induces spurious links and the consequent loss of information potentially encoded in those links. In the present work we used percolation on the trial-averaged TE network to prune the weaker links affected by the the inevitable

[1] For the complete description of the embedding methods and estimators for computation of H, which is beyond the scope of this study, see the works of Faes and colleagues[82,84,85] and references therein.

statistical noise introduced by the operation of averaging over trials. Since the networks analyzed here are weighted and directed it was necessary to apply the percolation analysis on the giant strongly connected components[95,96] and not on the connected components as usually done in the case of undirected graphs. Very briefly, a directed graph is strongly connected if every vertex is reachable from every other vertex following the direction of the links. Therefore, given an arbitrary directed graph a strongly connected component is a subgraph that is strongly connected.

To inspect the contribution of the hubs we measured the Vertex Degree (VD). VD is the number of links to a node i:

$$VD_{\mathrm{i}} = \sum_{j=1}^{N} a_{\mathrm{ij}} \ , \tag{4}$$

where a_{ij} is the generic entry of the adjacency matrix and N is the number of nodes. In directed networks one can distinguish between in-degree (the number on inward links) and out-degree (the number of outward links). We computed the probability distribution of both VDs (i.e. the in/out degree distribution) for each behavioural condition of each recording sessions for both animals. If the variance of the degree distribution is significantly larger than its mean, tails in the distribution arise and network hubs are identified (see Figure S1). Hubs are thus nodes that significantly exceed the average degree of the network[25]. Once hubs were identified, to further skim their contribution to the information exchange in the network we grouped them according to the behavioural condition in which they emerged. This selection procedure identified the classes. For example, if node i is an hub in Go trials but not in correct Stop trials, i would belong to the Go_drivers. In our framework, the TE values represented the strength of the connections between nodes and hence the classes are defined based on how much and during which behavioural condition nodes spread information trough the local PMd network.

Successively, in order to summarize in one number the total magnitude of information exchanged between the classes we constructed the following interaction measure I:

$$I_{\mathrm{ij}} = \rho \sum_{i}^{M} \sum_{j}^{M} \sum_{(m,n)} TE_{C_{\mathrm{i}}^{\mathrm{n}} \rightarrow C_{\mathrm{j}}^{\mathrm{n}}} \ , \tag{5}$$

where C is the module's class, M is the number of classes (M=4) and m and n run over the all possible combinations of nodes within each class. $\rho = \frac{1}{dim\,(C_{\mathrm{i}})}$ is a normalization factor that accounts for the heterogeneous number of nodes within each of the classes. Therefore, our $\mathrm{I_{ij}}$ is a normalized node strength computed on the graph formed by the 4 classes (i.e. in a weighted graph the strength of a node is the sum of the weights of the links connected to the node). Hence, the values and the directions of I values reflect the position in the hierarchy of the network communications for that class. All the interactions described by the empirical TE matrix were thus enclosed in a 4x4 matrix that represents a network of interactions in which now each node is a functional class. We computed I for each recording session and each behavioural condition and then we averaged over sessions for both animals. The uncertainty in estimating each element I_{ij} for each recording session was given by the standard error and thus the sessions-averaged element $\bar{I}_{\mathrm{ij}}$ (Figure 7 and supplementary Table S2) is estimated with an error obtained via the error propagation formula for the average of n measures.

4.6.1 A null model

To further assess the statistical significance of the results beyond what was already done by the randomization procedure included in the TE computations, we defined a null network model. As extensively detailed

in a recent work[97], the choice of a suitable null model remains a thorny issue in network and complex systems science. One common practice when analysing real-world networks such the one inspected in the present work, is try to identify properties that deviate from the null hypothesis, being likely that the deviations themselves encode information about the network functions. In this study, we drew our conclusions about the PMd information network in a completely data-driven fashion directly from neural activity, which thus constituted our only constraint. We thus tested whether the results were not attributable to the distribution of SA values and the dimensions of the corresponding time series. However, in the latter case our sample size was safely far from the limits of reliability reported in the literature for comparable systems[98,99] but we still used the null model as an additional control. Therefore we generated, for each single-trial of each behavioural condition and recording session, a synthetic pool of N time series with same length of the empirical ones (with N number of recording channels) by random sampling from the empirical SA distribution. We then computed the trial-averaged TE matrix for each synthetic pool. This situation is the most general since assumptions of any kind are made on the connectivity patterns and the weight distributions of the synthetic networks. We then compared the empirical measures with the ones obtained on the ensemble of 500 randomizations (see supplementary figure S5). In addition, we also compared our results with a well-established theoretical benchmark model naturally endowed with hubs: the Barabasi-Albert (BA) model[26]. Comparing again our results with the ones obtained on the ensemble of 500 randomizations of the BA model, we found, as expected, that our results were comparable with a hub-like network with the same number of nodes and density of links and not with the network derived from the SA distribution only.

Acknowledgements

Partially supported by Sapienza University (PH11715C823A9528), EU FP7 grant 'BrainLeap' (GA 306502), Fondazione G.B. Baroni (Borse di studio 2020), 'MUCCA' ERANET CHIST-ERA 2019 grant.

Author contributions statement

F.G, P.P., E.B. and S.F conceived the experiment, F.G. and P.P. conducted the experiment, G.B. analysed the results. All authors reviewed the manuscript.

Competing interests

The authors report no declarations of interest.

References

1. Hong, G. & Lieber, C. M. Novel electrode technologies for neural recordings. *Nat. Rev. Neurosci.* **20**, 330–345, 10.1038/s41583-019-0140-6 (2019).

2. Ferraina, S., Paré, M. & Wurtz, R. H. Comparison of cortico-cortical and cortico-collicular signals for the generation of saccadic eye movements. *J. Neurophysiol.* **87**, 845–858, 10.1152/jn.00317.2001 (2002).

3. Ferraina, S. *et al.* Early Coding of Visuomanual Coordination During Reaching in Parietal Area PEc. *J. Neurophysiol.* **85**, 462–467, 10.1152/jn.2001.85.1.462 (2001).

4. Wurtz, R. H., Sommer, M. A., Paré, M. & Ferraina, S. Signal transformations from cerebral cortex to superior colliculus for the generation of saccades. *Vis. Res.* **41**, 3399–3412, 10.1016/S0042-6989(01)00066-9 (2001).

5. Herreras, O. Local Field Potentials: Myths and Misunderstandings. *Front. Neural Circuits* **10**, 101, 10.3389/fncir.2016.00101 (2016).

6. Stark, E. & Abeles, M. Predicting movement from multiunit activity. *J. Neurosci.* **27**, 8387–8394, 10.1523/JNEUROSCI.1321-07.2007 (2007).

7. Trautmann, E. M. *et al.* Accurate Estimation of Neural Population Dynamics without Spike Sorting. *Neuron* **103**, 292–308, 10.1016/j.neuron.2019.05.003 (2019).

8. Mattia, M. *et al.* Heterogeneous attractor cell assemblies for motor planning in premotor cortex. *J. Neurosci.* **33**, 11155–11168, 10.1523/JNEUROSCI.4664-12.2013 (2013).

9. Riehle, A., Grün, S., Diesmann, M. & Aertsen, A. Spike synchronization and rate modulation differentially involved in motor cortical function. *Science* **278**, 1950–1953, 10.1126/science.278.5345.1950 (1997).

10. Georgopoulos, A. P., Schwartz, A. B. & Kettner, R. E. Neuronal population coding of movement direction. *Science* **233**, 1416–1419, 10.1126/science.3749885 (1986).

11. Munoz, D. P. & Wurtz, R. H. Fixation cells in monkey superior colliculus I. Characteristics of cell discharge. *J. Neurophysiol.* **70**, 559–575, 10.1152/jn.1993.70.2.559 (1993).

12. Schall, J. D. Visuomotor Functions in the Frontal Lobe. *Annu. Rev. Vis. Sci.* **1**, 469–498, 10.1146/annurev-vision-082114-035317 (2015).

13. Vaadia, E., Kurata, K. & Wise, S. P. Neuronal activity preceding directional and nondirectional cues in the premotor cortex of rhesus monkeys. *Somatosens. & Mot. Res.* **6**, 207–230, 10.3109/08990228809144674 (1988).

14. Fetz, E. E. Are movement parameters recognizably coded in the activity of single neurons? *Behav. Brain Sci.* **15**, 679–690, 10.1017/S0140525X00072599 (1992).

15. Vyas, S., Golub, M. D., Sussillo, D. & Shenoy, K. V. Computation through Neural Population Dynamics, 10.1146/annurev-neuro-092619-094115 (2020).

16. Bardella, G., Pani, P., Brunamonti, E., Giarrocco, F. & Ferraina, S. The small scale functional topology of movement control: Hierarchical organization of local activity anticipates movement generation in the premotor cortex of primates. *NeuroImage* **207**, 10.1016/j.neuroimage.2019.116354 (2020).

17. Giarrocco, F. *et al.* Neuronal dynamics of signal selective motor plan cancellation in the macaque dorsal premotor cortex. *Cortex* **135**, 326–340, 10.1016/j.cortex.2020.09.032 (2021).

18. Hubel, D. H. & Wiesel, T. N. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The J. Physiol.* **160**, 106–154, 10.1113/jphysiol.1962.sp006837 (1962).

19. Wise, S. P., Boussaoud, D., Johnson, P. B. & Caminiti, R. Premotor and parietal cortex: Corticocortical connectivity and combinatorial computations, 10.1146/annurev.neuro.20.1.25 (1997).

20. Cisek, P. & Kalaska, J. F. Neural mechanisms for interacting with a world full of action choices, 10.1146/annurev.neuro.051508.135409 (2010).

21. Shenoy, K. V., Sahani, M. & Churchland, M. M. Cortical Control of Arm Movements: A Dynamical Systems Perspective. *Annu. Rev. Neurosci.* **36**, 337–359, 10.1146/annurev-neuro-062111-150509 (2013).

22. Milekovic, T., Truccolo, W., Grün, S., Riehle, A. & Brochier, T. Local field potentials in primate motor cortex encode grasp kinetic parameters. *NeuroImage* **114**, 338–355, 10.1016/J.NEUROIMAGE.2015.04.008 (2015).

23. Georgopoulos, A. P., Kalaska, J. F., Caminiti, R. & Massey, J. T. On the relations between the direction of two-dimensional arm movements and cell discharge in primate motor cortex. *The J. neuroscience : official journal Soc. for Neurosci.* **2**, 1527–1537, 10.1523/JNEUROSCI.02-11-01527.1982 (1982).

24. Caminiti, R., Johnson, P. B., Galli, C., Ferraina, S. & Burnod, Y. Making arm movements within different parts of space: the premotor and motor cortical representation of a coordinate system for reaching to visual targets. *J. Neurosci.* **11**, 1182–1197, 10.1523/JNEUROSCI.11-05-01182.1991 (1991).

25. Albert, R. & Barabási, A.-L. Statistical mechanics of complex networks. *Rev. Mod. Phys.* **74**, 47–97, 10.1103/RevModPhys.74.47 (2002).

26. Barabási, A. L. Network science, 10.1098/rsta.2012.0375 (2013).

27. Shimono, M. & Beggs, J. M. Functional clusters, hubs, and communities in the cortical microconnectome. *Cereb. Cortex* **25**, 3743–3757, 10.1093/cercor/bhu252 (2015).

28. Buehlmann, A. & Deco, G. Optimal information transfer in the cortex through synchronization. *PLoS Comput. Biol.* **6**, 1000934, 10.1371/journal.pcbi.1000934 (2010).

29. Orlandi, J. G., Stetter, O., Soriano, J., Geisel, T. & Battaglia, D. Transfer entropy reconstruction and labeling of neuronal connections from simulated calcium imaging. *PLoS ONE* **9**, 98842, 10.1371/journal.pone.0098842 (2014).

30. Ito, S. *et al.* Extending Transfer Entropy Improves Identification of Effective Connectivity in a Spiking Cortical Network Model. *PLoS ONE* **6**, e27431, 10.1371/journal.pone.0027431 (2011).

31. Nigam, S. *et al.* Rich-club organization in effective connectivity among cortical neurons. *J. Neurosci.* **36**, 655–669, 10.1523/JNEUROSCI.2177-15.2016 (2016).

32. Timme, N. M. *et al.* High-Degree Neurons Feed Cortical Computations. *PLOS Comput. Biol.* **12**, e1004858, 10.1371/journal.pcbi.1004858 (2016).

33. Gerhard, F., Pipa, G., Lima, B., Neuenschwander, S. & Gerstner, W. Extraction of Network Topology From Multi-Electrode Recordings: Is there a Small-World Effect? *Front. Comput. Neurosci.* **5**, 4, 10.3389/fncom.2011.00004 (2011).

34. Besserve, M., Lowe, S. C., Logothetis, N. K., Schölkopf, B. & Panzeri, S. Shifts of Gamma Phase across Primary Visual Cortical Sites Reflect Dynamic Stimulus-Modulated Information Transfer. *PLoS Biol.* **13**, e1002257, 10.1371/journal.pbio.1002257 (2015).

35. Honey, C. J., Kötter, R., Breakspear, M. & Sporns, O. Network structure of cerebral cortex shapes functional connectivity on multiple time scales. *Proc. Natl. Acad. Sci. United States Am.* **104**, 10240–10245, 10.1073/pnas.0701519104 (2007).

36. Jahfari, S. *et al.* Effective connectivity reveals important roles for both the hyperdirect (fronto-subthalamic) and the indirect (fronto-striatal-pallidal) fronto-basal ganglia pathways during response inhibition. *J. Neurosci.* **31**, 6891–6899, 10.1523/JNEUROSCI.5253-10.2011 (2011).

37. Boucher, L., Palmeri, T. J., Logan, G. D. & Schall, J. D. Inhibitory control in mind and brain: An interactive race model of countermanding saccades. *Psychol. Rev.* **114**, 376–397, 10.1037/0033-295X.114.2.376 (2007).

38. Lo, C. C., Boucher, L., Paré, M., Schall, J. D. & Wang, X. J. Proactive inhibitory control and attractor dynamics in countermanding action: A spiking neural circuit model. *J. Neurosci.* **29**, 9059–9071, 10.1523/JNEUROSCI.6164-08.2009 (2009).

39. Churchland, M. M. *et al.* Neural population dynamics during reaching. *Nature* **487**, 51–56, 10.1038/nature11129 (2012).

40. Chandrasekaran, C., Peixoto, D., Newsome, W. T. & Shenoy, K. V. Laminar differences in decision-related neural activity in dorsal premotor cortex. *Nat. Commun.* **8**, 1–16, 10.1038/s41467-017-00715-0 (2017).

41. Kaufman, M. T. *et al.* The largest response component in the motor cortex reflects movement timing but not movement type. *eNeuro* **3**, 85–101, 10.1523/ENEURO.0085-16.2016 (2016).

42. Clawson, W. *et al.* Computing hubs in the hippocampus and cortex. *Sci. Adv.* **5**, eaax4843, 10.1126/sciadv.aax4843 (2019).

43. Schroeder, K. E. *et al.* Disruption of corticocortical information transfer during ketamine anesthesia in the primate brain. *NeuroImage* **134**, 459–465, 10.1016/j.neuroimage.2016.04.039 (2016).

44. Marcos, E. *et al.* Neural variability in premotor cortex is modulated by trial history and predicts behavioral performance. *Neuron* **78**, 249–255, 10.1016/j.neuron.2013.02.006 (2013).

45. Verbruggen, F. & Logan, G. D. Response inhibition in the stop-signal paradigm, 10.1016/j.tics.2008.07.005 (2008).

46. Schall, J. D., Palmeri, T. J. & Logan, G. D. Models of inhibitory control. *Philos. Transactions Royal Soc. B: Biol. Sci.* **372**, 20160193, 10.1098/rstb.2016.0193 (2017).

47. Wei, W., Rubin, J. E. & Wang, X. J. Role of the indirect pathway of the basal ganglia in perceptual decision making. *J. Neurosci.* **35**, 4052–4064, 10.1523/JNEUROSCI.3611-14.2015 (2015).

48. Gu, Y., Qi, Y. & Gong, P. Rich-club connectivity, diverse population coupling, and dynamical activity patterns emerging from local cortical circuits. *PLoS Comput. Biol.* **15**, e1006902, 10.1371/journal.pcbi.1006902 (2019).

49. Perin, R., Berger, T. K. & Markram, H. A synaptic organizing principle for cortical neuronal groups. *Proc. Natl. Acad. Sci. United States Am.* **108**, 5419–5424, 10.1073/pnas.1016051108 (2011).

50. Timme, N. *et al.* Multiplex Networks of Cortical and Hippocampal Neurons Revealed at Different Timescales. *PLoS ONE* **9**, e115764, 10.1371/journal.pone.0115764 (2014).

51. Schroeter, M. S., Charlesworth, P., Kitzbichler, M. G., Paulsen, O. & Bullmore, E. T. Emergence of rich-club topology and coordinated dynamics in development of hippocampal functional networks In vitro. *J. Neurosci.* **35**, 5459–5470, 10.1523/JNEUROSCI.4259-14.2015 (2015).

52. Gal, E. *et al.* Rich cell-type-specific network topology in neocortical microcircuitry. *Nat. Neurosci.* **20**, 1004–1013, 10.1038/nn.4576 (2017).

53. Dann, B., Michaels, J. A., Schaffelhofer, S. & Scherberger, H. Uniting functional network topology and oscillations in the fronto-parietal single unit network of behaving primates. *eLife* **5**, 10.7554/eLife.15719 (2016).

54. Hilgetag, C., O'Neill, M. A. & Young, M. P. Hierarchical organization of macaque and cat cortical sensory systems explored with a novel network processor. *Philos. Transactions Royal Soc. London. Ser. B: Biol. Sci.* **355**, 71–89, 10.1098/rstb.2000.0550 (2000).

55. Hilgetag, C. C. & Kaiser, M. Clustered organization of cortical connectivity, 10.1385/NI:2:3:353 (2004).

56. Felleman, D. J., Felleman, D. J. & Van Essen, D. C. Distributed hierarchical processing in the primate cerebral cortex. *CEREB CORTEX* 1–47 (1991).

57. Sporns, O., Tononi, G. & Kötter, R. The Human Connectome: A Structural Description of the Human Brain. *PLoS Comput. Biol.* **1**, e42, 10.1371/journal.pcbi.0010042 (2005).

58. Kaiser, M. Hierarchy and dynamics of neural networks. *Front. Neuroinformatics* **4**, 112, 10.3389/fninf.2010.00112 (2010).

59. Zamora-López. Cortical hubs form a module for multisensory integration on top of the hierarchy of cortical networks. *Front. Neuroinformatics* **4**, 1, 10.3389/neuro.11.001.2010 (2010).

60. Zeki, S. & Shipp, S. The functional logic of cortical connections, 10.1038/335311a0 (1988).

61. Tononi, G., Sporns, O. & Edelman, G. M. A measure for brain complexity: Relating functional segregation and integration in the nervous system. *Proc. Natl. Acad. Sci. United States Am.* **91**, 5033–5037, 10.1073/pnas.91.11.5033 (1994).

62. Sporns, O., Honey, C. J. & Kötter, R. Identification and Classification of Hubs in Brain Networks. *PLoS ONE* **2**, e1049, 10.1371/journal.pone.0001049 (2007).

63. Watanabe, T. *et al.* Energy landscapes of resting-state brain networks. *Front. Neuroinformatics* **8**, 12, 10.3389/fninf.2014.00012 (2014).

64. Kaiser, M., Görner, M. & Hilgetag, C. C. Criticality of spreading dynamics in hierarchical cluster networks without inhibition. *New J. Phys.* **9**, 110, 10.1088/1367-2630/9/5/110 (2007).

65. Bardella, G., Bifone, A., Gabrielli, A., Gozzi, A. & Squartini, T. Hierarchical organization of functional connectivity in the mouse brain: A complex network approach. *Sci. Reports* **6**, 1–11, 10.1038/srep32060 (2016).

66. Hilgetag, C. C. & Goulas, A. 'Hierarchy' in the organization of brain networks. *Philos. Transactions Royal Soc. B: Biol. Sci.* **375**, 20190319, 10.1098/rstb.2019.0319 (2020).

67. Fox, M. D. *et al.* The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proc. Natl. Acad. Sci. United States Am.* **102**, 9673–9678, 10.1073/pnas.0504136102 (2005).

68. He, B. J. Scale-free properties of the functional magnetic resonance imaging signal during rest and task. *J. Neurosci.* **31**, 13786–13795, 10.1523/JNEUROSCI.2111-11.2011 (2011).

69. Zalesky, A., Fornito, A., Cocchi, L., Gollo, L. L. & Breakspear, M. Time-resolved resting-state brain networks. *Proc. Natl. Acad. Sci. United States Am.* **111**, 10341–10346, 10.1073/pnas.1400181111 (2014).

70. Shine, J. M. *et al.* The Dynamics of Functional Brain Networks: Integrated Network States during Cognitive Task Performance. *Neuron* **92**, 544–554, 10.1016/j.neuron.2016.09.018 (2016).

71. Vidaurre, D., Smith, S. M. & Woolrich, M. W. Brain network dynamics are hierarchically organized in time. *Proc. Natl. Acad. Sci. United States Am.* **114**, 12827–12832, 10.1073/pnas.1705120114 (2017).

72. Morcos, A. S. & Harvey, C. D. History-dependent variability in population dynamics during evidence accumulation in cortex. *Nat. Neurosci.* **19**, 1672–1681, 10.1038/nn.4403 (2016).

73. Mirabella, G., Pani, P. & Ferraina, S. Neural correlates of cognitive control of reaching movements in the dorsal premotor cortex of rhesus monkeys. *J. Neurophysiol.* **106**, 1454–1466, 10.1152/jn.00995.2010 (2011).

74. Logan, G. D. & Cowan, W. B. On the ability to inhibit thought and action: A theory of an act of control. *Psychol. Rev.* **91**, 295–327, 10.1037//0033-295x.91.3.295 (1984).

75. Band, G. P., van der Molen, M. W. & Logan, G. D. Horse-race model simulations of the stop-signal procedure. *Acta Psychol.* **112**, 105–142, 10.1016/S0001-6918(02)00079-3 (2003).

76. Mattia, M. & Del Giudice, P. Population dynamics of interacting spiking neurons. *Phys. Rev. E - Stat. Physics, Plasmas, Fluids, Relat. Interdiscip. Top.* **66**, 19, 10.1103/PhysRevE.66.051917 (2002).

77. Xiong, W., Faes, L. & Ivanov, P. C. Entropy measures, entropy estimators, and their performance in quantifying complex dynamics: Effects of artifacts, nonstationarity, and long-range correlations. *Phys. Rev. E* **95**, 1–37, 10.1103/PhysRevE.95.062114 (2017).

78. Wollstadt, P., Martínez-Zarzuela, M., Vicente, R., Díaz-Pernas, F. J. & Wibral, M. Efficient transfer entropy analysis of non-stationary neural time series. *PLoS ONE* **9**, 102833, 10.1371/journal.pone.0102833 (2014).

79. Vicente, R., Wibral, M., Lindner, M. & Pipa, G. Transfer entropy-a model-free measure of effective connectivity for the neurosciences. *J. Comput. Neurosci.* **30**, 45–67, 10.1007/s10827-010-0262-3 (2011).

80. Wibral, M. *et al.* Transfer entropy in magnetoencephalographic data: Quantifying information flow in cortical and cerebellar networks. *Prog. Biophys. Mol. Biol.* **105**, 80–97, 10.1016/j.pbiomolbio.2010.11.006 (2011).

81. Montalto, A., Faes, L. & Marinazzo, D. MuTE: A MATLAB toolbox to compare established and novel estimators of the multivariate transfer entropy. *PLoS ONE* **9**, e109462, 10.1371/journal.pone.0109462 (2014).

82. Faes, L., Kugiumtzis, D., Nollo, G., Jurysta, F. & Marinazzo, D. Estimating the decomposition of predictive information in multivariate systems. *PHYSICAL REVIEW E* **91**, 32904, 10.1103/PhysRevE.91.032904 (2015).

83. Shannon, C. E. A Mathematical Theory of Communication. Tech. Rep.

84. Faes, L., Marinazzo, D., Montalto, A. & Nollo, G. Lag-specific transfer entropy as a tool to assess cardiovascular and cardiorespiratory information transfer. *IEEE Transactions on Biomed. Eng.* **61**, 2556–2568, 10.1109/TBME.2014.2323131 (2014).

85. Faes, L., Nollo, G. & Porta, A. Information-based detection of nonlinear Granger causality in multivariate processes via a nonuniform embedding technique. *Phys. Rev. E - Stat. Nonlinear, Soft Matter Phys.* **83**, 051112, 10.1103/PhysRevE.83.051112 (2011).

86. Callaway, D. S., Newman, M. E., Strogatz, S. H. & Watts, D. J. Network Robustness and Fragility: Percolation on Random Graphs. *Phys. Rev. Lett.* **85**, 5468, 10.1103/PhysRevLett.85.5468 (2000).

87. Gallos, L. K., Makse, H. A. & Sigman, M. A small world of weak ties provides optimal global integration of self-similar modules in functional brain networks. *Proc. Natl. Acad. Sci. United States Am.* **109**, 2825–2830, 10.1073/PNAS.1106612109/-/DCSUPPLEMENTAL/PNAS.1106612109{_}SI.PDF (2012).

88. Nicolini, C., Bordier, C. & Bifone, A. Community detection in weighted brain connectivity networks beyond the resolution limit. *NeuroImage* **146**, 28–39, 10.1016/j.neuroimage.2016.11.026 (2017).

89. Nicolini, C. & Bifone, A. Modular structure of brain functional networks: breaking the resolution limit by Surprise. *Sci. Reports* **6**, 10.1038/SREP19250 (2016).

90. Vlasov, V. & Bifone, A. Hub-driven remote synchronization in brain networks. *Sci. Reports 2017 7:1* **7**, 1–11, 10.1038/s41598-017-09887-7 (2017).

91. Mastrandrea, R. *et al.* Organization and hierarchy of the human functional brain network lead to a chain-like core. *Sci. Reports 2017 7:1* **7**, 1–13, 10.1038/s41598-017-04716-3 (2017).

92. Bordier, C., Nicolini, C. & Bifone, A. Graph analysis and modularity of brain functional connectivity networks: Searching for the optimal threshold. *Front. Neurosci.* **11**, 441, 10.3389/FNINS.2017.00441/BIBTEX (2017).

93. Bordier, C., Nicolini, C., Forcellini, G. & Bifone, A. Disrupted modular organization of primary sensory brain areas in schizophrenia. *NeuroImage: Clin.* **18**, 682–693, 10.1016/j.nicl.2018.02.035 (2018).

94. Del Ferraro, G. *et al.* Finding influential nodes for integration in brain networks using optimal percolation theory. *Nat. Commun. 2018 9:1* **9**, 1–12, 10.1038/s41467-018-04718-3 (2018).

95. Graf, A. On the Strongly Connected Components of Random Directed Graphs with Given Degree Sequences. Tech. Rep., University of Waterloo (2016).

96. Squires, S. *et al.* Weakly explosive percolation in directed networks. *PHYSICAL REVIEW E* **87**, 52127, 10.1103/PhysRevE.87.052127 (2013).

97. Cimini, G. *et al.* The statistical physics of real-world networks. *Nat. Rev. Phys.* **1**, 58–71, 10.1038/s42254-018-0002-6 (2019).

98. Ramos, A. M. T., Macau, E. E. N., Lopes, A. M. & Machado, J. A. T. Minimum Sample Size for Reliable Causal Inference Using Transfer Entropy. *Entropy 2017, Vol. 19, Page 150* **19**, 150, 10.3390/E19040150 (2017).

99. Timme, N. M. & Lapish, C. A Tutorial for Information Theory in Neuroscience. *eNeuro* **5**, 10.1523/ENEURO.0052-18.2018 (2018).