

RESEARCH

Open Access



# Behavioural relevance of redundant and synergistic stimulus information between functionally connected neurons in mouse auditory cortex

Loren Koçillari<sup>1,2,3\*</sup>, Marco Celotto<sup>1,2,4</sup>, Nikolas A. Francis<sup>5</sup>, Shoutik Mukherjee<sup>6</sup>, Behtash Babadi<sup>6</sup>, Patrick O. Kanold<sup>7</sup> and Stefano Panzeri<sup>2\*</sup>

## Abstract

Measures of functional connectivity have played a central role in advancing our understanding of how information is transmitted and processed within the brain. Traditionally, these studies have focused on identifying redundant functional connectivity, which involves determining when activity is similar across different sites or neurons. However, recent research has highlighted the importance of also identifying synergistic connectivity—that is, connectivity that gives rise to information not contained in either site or neuron alone. Here, we measured redundant and synergistic functional connectivity between neurons in the mouse primary auditory cortex during a sound discrimination task. Specifically, we measured directed functional connectivity between neurons simultaneously recorded with calcium imaging. We used Granger Causality as a functional connectivity measure. We then used Partial Information Decomposition to quantify the amount of redundant and synergistic information about the presented sound that is carried by functionally connected or functionally unconnected pairs of neurons. We found that functionally connected pairs present proportionally more redundant information and proportionally less synergistic information about sound than unconnected pairs, suggesting that their functional connectivity is primarily redundant. Further, synergy and redundancy coexisted both when mice made correct or incorrect perceptual discriminations. However, redundancy was much higher (both in absolute terms and in proportion to the total information available in neuron pairs) in correct behavioural choices compared to incorrect ones, whereas synergy was higher in absolute terms but lower in relative terms in correct than in incorrect behavioural choices. Moreover, the proportion of redundancy reliably predicted perceptual discriminations, with the proportion of synergy adding no extra predictive power. These results suggest a crucial contribution of redundancy to correct perceptual discriminations, possibly due to the advantage it offers for information propagation, and also suggest a role of synergy in enhancing information level during correct discriminations.

**Keywords** Functional connectivity, Redundancy, Synergy, Auditory perception

\*Correspondence:

Loren Koçillari  
loren.kocillari@iit.it  
Stefano Panzeri  
s.panzeri@uke.de

Full list of author information is available at the end of the article



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## 1 Introduction

Functional connectivity (FC) has emerged as a mainstream concept and a fundamental tool for understanding how brain networks process and communicate information, and how functional interactions between networks or between neurons shape the dynamics and function of the brain [1–10]. Traditional measures of FC have mainly focused on redundant connectivity, by measuring (for example, through linear cross-correlations) the similarity of activity between different sites. However, recent studies have begun to highlight the importance of another notion of FC: synergistic connectivity [11–15]. This notion of connectivity focuses on how variations of the interaction between activity at different sites or between activity of different neurons create information that is not present at each site or in each neuron alone [4, 14–16]. Whilst the presence and merits of redundant connectivity have been extensively documented [1, 3, 17, 18], it remains unclear whether synergistic interactions are prominent and how they contribute to cognitive function.

Correlated activity is present in multiple spatial scales, from brain areas to local networks. Consequently, an additional question pertains to the spatial scale at which both redundant and synergistic interactions are expressed. Most previous studies of FC investigated it at a coarse scale, such as that obtained with non-invasive measures of neural activity, such as fMRI or EEG, that do not have single-neuron resolution [19–23]. However, the organization of FC at the finer spatial scale of population recordings with single-neuron resolution is less understood, and its relationship to redundancy or synergy of information encoding at this finer scale has been considered only seldom [24].

In this study, we address some of these open questions regarding synergistic and redundant FC. First, we address their relationship with respect to a widely used directed FC measure, Granger Causality (GC) [25, 26], between the activities of different neurons. This measure of FC is interesting because, unlike simple measures of FC based on cross-correlation, it considers not only the similarity of activity but also the strength and directionality of information transmission. GC, as well as other measures implementing the Wiener–Granger Causality principle [27], can, in principle, capture redundant FC because the process of transmission entails sharing of information between the sending and the receiving site [28, 29]. However, it can also correspond to synergistic FC. For example, if transmission varies across sensory stimuli, FC can create sensory information not available in each site individually. Second, we use precise information-theoretic measures to quantify redundancy and synergy related to the encoding of behaviourally relevant sensory variables

(in this case, features of auditory stimuli). These measures, based on the theory of Partial Information Decomposition (PID) [30, 31], have the advantage of separating redundancy from synergy, something that simpler measures [32] used in recent studies [33, 34] cannot achieve. Third, we study synergistic and redundant stimulus information with single-neuron resolution, using the primary auditory cortex (A1) of the mouse brain as an experimental model. Fourth, we explore the potential impact of synergy and redundancy on sensory processing by studying how they vary between cases of correct and incorrect perceptual discrimination.

Part of this work has been presented at the 16th International Conference of Brain Informatics and published as a conference paper [35].

## 2 Experimental task and single-neuron stimulus information

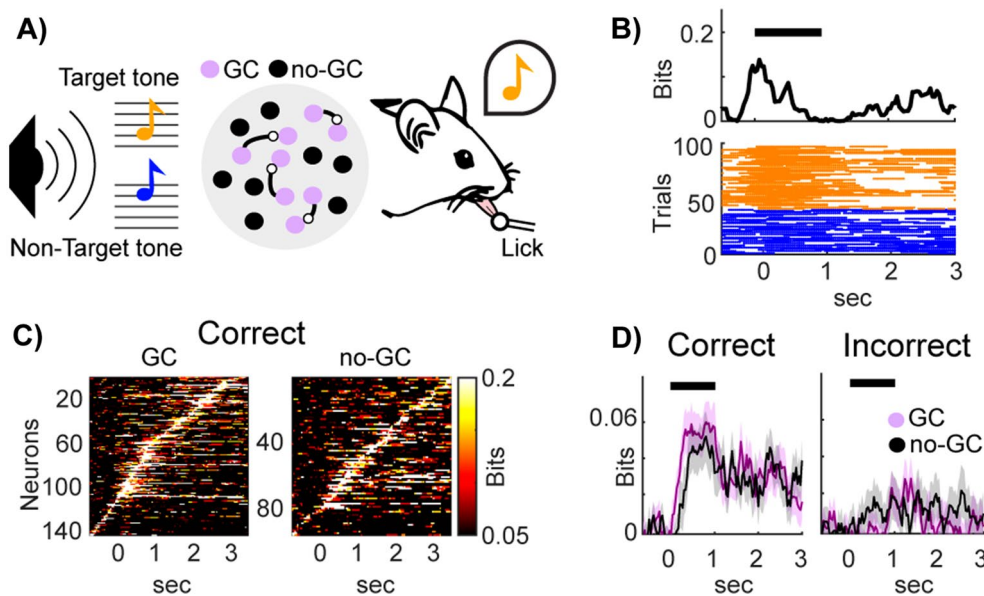
To investigate the relationship between FC and the presence of synergistic and redundant information with single-neuron resolution, we focused on the activity of the mouse primary auditory cortex during a sound discrimination task. We reanalysed a previously published dataset [34] in which the activity of several tens to a few hundreds of neurons was recorded simultaneously using *in vivo* two photon calcium imaging from A1 L2/3 neurons in transgenic mice during a pure-tone discrimination task (Fig. 1A).

The experimental task was structured as follows. After a pre-stimulus interval of 1 s, head-fixed mice were exposed to either a low-frequency (7 or 9.9 kHz) or a high-frequency (14 or 19.8 kHz) tone for a period of 1 s. Mice were trained to report their perception of the sound stimulus by their behavioural choice, which consisted of licking a waterspout in the post-stimulus interval (0.5–3 s from stimulus onset) after hearing a low-frequency tone (target tones) and holding still after hearing high-frequency tones (non-target tones). Calcium imaging was used to continuously acquire the fluorescence signals from individual A1 L2/3 neurons during the task with a sampling frequency of 30 Hz.

We used Shannon mutual information [36, 37] to compute the stimulus information carried by each neuron about the stimulus category (low- vs high-frequency tones) in each imaging time frame (Fig. 1B, top plot). Stimulus information is defined as follows:

$$SI(S; R_i) = \sum_{s \in S, r_i \in R} p(s, r_i) \log_2 \frac{p(s, r_i)}{p(s)p(r_i)}, \quad (1)$$

where  $i$  indexes the neurons and  $p(s, r_i)$  denotes the joint probability of observing in a given trial the activity  $r_i$  of neuron  $i$  and the value  $s$  of the stimulus variable  $S$ .



**Fig. 1** Stimulus information in mouse auditory cortex during a tone discrimination task. **A** Mice performed a go/no-go tone discrimination task whilst the activity of A1 L2/3 neurons were recorded with two-photon calcium imaging. In response to a target tone (low-frequency, in orange) mice had to lick a waterspout and not to lick for non-target tones (high-frequency, in blue). Granger causality analysis revealed sparsely connected networks of cells in A1 L2/3 [34]. We classified neurons as GC (purple) and no-GC (black) depending on whether they formed a GC link. **B** Top: Example of the stimulus information time-course for a single neuron. We computed the time-resolved stimulus information as the mutual information between the auditory stimuli (low-/high-frequency tones) and the spiking activity across trials. Bottom: The traces of the deconvolved spiking activity in each trial, colour-coded based on the tone presented in each trial. The time axes are referenced based on the stimulus onset. **C** Stimulus information time-course for GC neurons (left map) and no-GC neurons (right map) in correct trials only. We then sorted the peaks of stimulus information for each neuron to tile the trial time. **D** Stimulus information for GC (in purple) and no-GC (black) neurons computed separately in trials with correct and incorrect choices. Full lines denote the mean across neurons, whilst the shaded areas denote the SEM across neurons

$p(r_i) = \sum_s p(s, r_i)$  and  $p(s) = \sum_{r_i} p(s, r_i)$  are the marginal probabilities.

The activity  $r_i$  of neuron  $i$  was inferred following the same approach described in [34]. In brief, we first deconvolved the single-trial calcium fluorescence traces of each neuron to infer the spiking activity (Fig. 1B, bottom). We then aligned neural activity of each trial to the stimulus onset. We used a sliding window approach (windows of 10 time frames with time-steps of 1 time frame) to binarize the deconvolved spiking activity of each window into 0 and 1, where 1 denotes spiking activity higher than 0. We then computed the time-resolved stimulus information on these binarized neural responses using the probabilities  $p(s, r_i)$  obtained empirically from the data and plugging them into the information theoretic equations [38]. Finally, we subtracted the average stimulus information computed in the pre-stimulus interval from the stimulus information time-courses, which enabled us to correct for the systematic error (or bias) in the information estimate due to the limited number of trials [39].

Following our previous study [34], we first analysed the entire dataset (2792 neurons recorded from 34 sessions) to identify those neurons that carried significant

task-related information. Neurons were defined as carrying task-related information if they carried statistically significant stimulus information (defined as in Eq. (1) above), significant choice information (defined as in Eq. (1) above but replacing the stimulus presented in the given trial with the behavioural choice of the animal in the trial), and intersection information. In brief, intersection information uses the mathematical framework of PID [30, 40] to quantify the amount of sensory information encoded in neural activity that is used to inform the behavioural choice [41]. It satisfies a number of information theoretic properties that would be expected of such a measure, including being upper bounded by the stimulus information encoded in neural activity, the choice information encoded in neural activity and by the mutual information between stimulus and choice [40].

The statistical significance of each information measure was computed using a non-parametric permutation test at  $p < 0.1$  on the information time-courses. We generated a null hypothesis distribution by randomly shuffling the associations between stimuli and neural responses, or between choices and neural responses, across trials at each time point. For each random permutation, we

selected the highest information value across all time windows. We then calculated the  $p$ -values based on how often the peaks of information in the shuffled dataset exceeded those in the actual dataset. By considering the highest information value as the summary statistic for each trial, the  $p$ -values obtained as such are already adjusted for multiple comparisons across time. The requirement of all three non-independent tests being satisfied simultaneously was empirically estimated, resulting in a false discovery rate of 1% [34].

We found a subset of 475/2790 neurons that transiently and sequentially carried significant task-relevant information [34]. Using methods described in [26, 34] we next performed GC analysis on this subset of neurons. We selected 20 neurons per session, with peak intersection information exhibiting the shortest latencies. If more than 20 such neurons were present, the 20 with the shortest latencies were selected. We focused our analyses on 12 out of 34 sessions that had at least 20 neurons with significant intersection information. We found that these neurons formed sparse functional networks that transmitted redundant task-relevant information across the trial time (Fig. 1A) [34]. Of these 240 neurons, 144 formed GC connections with at least another neuron (out of the network of 20 neurons that has significant intersection information in the same session) and were termed GC neurons hereafter. The remaining 96 neurons, which did not form GC connections with any other neuron, were termed no-GC neurons hereafter.

We used information-theoretic measures to quantify the stimulus information dynamics of individual neurons. We first considered information in trials in which the mouse made correct perceptual discriminations. The stimulus information time-courses, plotted in (Fig. 1C) after sorting neurons by their peak information timing, showed sequential information coding across the population in both GC and no-GC neurons. At peak, neurons had similar amounts of information in both populations, with the main difference being that GC neurons exhibited the peak information earlier in the trial (during stimulus presentation), whilst no-GC neurons carried information later in the trial (after stimulus presentation) (Fig. 1C). The sequential nature of their activation suggests that information is represented throughout the trial only at the population level, motivating our later information analyses at the neural population level.

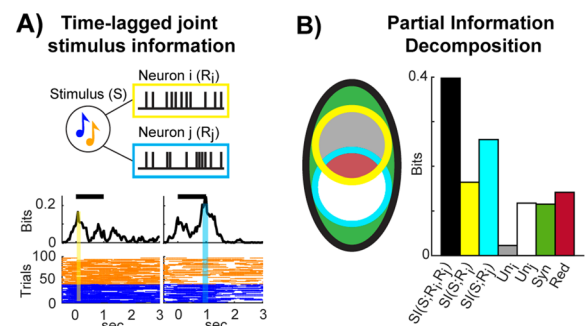
To investigate what aspects of neural activity may be a key for correct perceptual judgements, we assessed how information about the auditory stimulus category was encoded in trials in which the animal judged the sound stimulus either correctly or incorrectly. The average stimulus information across all neurons is reported in Fig. 1D. Importantly, we found that the stimulus information was

lower in incorrect than in correct trials for both GC and no-GC neurons across the entire trial time, suggesting that the stimulus information is used for the behavioural choice. Importantly, all information quantities computed during correct discriminations were calculated on random subsets of correct trials with the same size as the number of incorrect trials in the same session. Due to this balanced sub-sampling strategy, we were able to make a fair comparison of the amount of information encoded in correct and incorrect trials and control for potential systematic errors due to limited-sampling bias [39].

### 3 Emergent properties of population codes in auditory cortex during correct and incorrect behaviour

We next asked how correct and incorrect behaviour relates to the emergent properties of population codes. This required computing stimulus information from more than 1 neuron.

As in our previous study [34], we estimated the total stimulus information that was jointly carried by pairs of neurons following a time-lagged approach (Fig. 2A). We first identified for each neuron the peak time of task-related information, i.e., the time frame when



**Fig. 2** Partial information decomposition of the time-lagged joint stimulus information. **A** Top: Schematic of the computation of the time-lagged joint stimulus information, defined as the mutual information that the neural responses of two neurons at their information time peaks (yellow and cyan vertical bars in the plots on the bottom) jointly carry about the stimulus category. Bottom: The stimulus information time-course (as function of peristimulus time) of two example neurons as a black line, and the deconvolved spiking activity in each trial as orange lines for the low-frequency tones and as a blue line for the high-frequency tones. **B** Left: Sketch as Venn diagram of the decomposition of the joint stimulus information into the non-negative components of synergistic (green area), redundant (red area), and unique (grey and white areas) information. The stimulus information carried by individual neurons (in yellow and cyan) is the sum of the redundant and unique information components. The bar plot on the right shows the amount of the joint stimulus information carried by the example pair of neurons in **A** (black bar), the stimulus information carried by each neuron individually (yellow and cyan bars), the unique (grey and white bars), synergistic (green bar), and redundant (red bar) information

intersection information time-courses peaked. We then computed the time-lagged stimulus information carried jointly by the activity of each pair of neurons as follows:

$$SI(S; R_i, R_j) = \sum_{s \in S, r_i \in R_i, r_j \in R_j} p(s, r_i, r_j) \log_2 \frac{p(s, r_i, r_j)}{p(s)p(r_i, r_j)}, \quad (2)$$

where  $p(s, r_i, r_j)$  denotes the probability of simultaneously observing in the same trial the value  $s$  of the stimulus category and the joint neural responses  $r_i$  and  $r_j$  of neurons  $i$  and  $j$  measured at their respective peaks of task-related information.

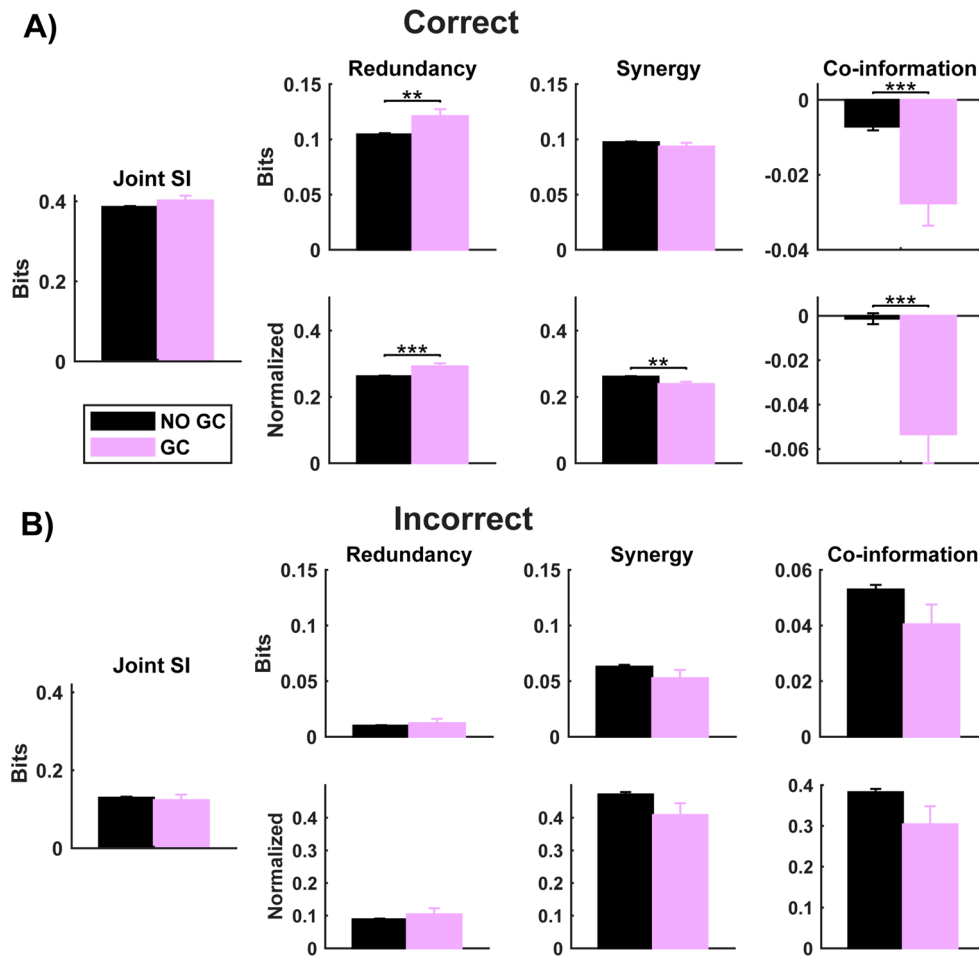
First, following our previous work [34], we investigated the nature of redundant and synergistic interactions in pairs of neurons by computing the so-called co-information [42], defined as the difference between the total

stimulus information that was jointly carried by both neurons (Eq. (2)) and the sum of stimulus information carried by each neuron individually (Eq. (1)):

$$CoInfo(S; R_i; R_j) = SI(S; R_i, R_j) - (SI(S; R_i) + SI(S; R_j)) \quad (3)$$

A positive value of  $CoInfo(S; R_i; R_j)$  implies that the pair of neurons carries more information than the sum of their individual information and can thus be interpreted as predominant synergy. Similarly, a negative value can be interpreted as predominant redundancy.

As in our previous study [34], on average across pairs of neurons we found negative co-information (indicating predominance of redundancy) in correct trials and positive co-information (indicating predominance of synergy) in incorrect trials (Fig. 3A).



**Fig. 3** Redundancy and synergy in auditory cortex. **A** Redundancy and synergy computed in correct trials. From left to right: time-lagged joint stimulus information (Joint SI), redundancy, synergy, and co-information for GC-connected (purple) and GC-unconnected pairs of neurons (black). For synergy, redundancy and co-information, the top plots show values in bits and the bottom plots show values normalized by the joint stimulus information. **B** as in panel **A**, but in the case of incorrect trials. Bar plots show mean  $\pm$  SEM across pairs. Statistics were made with a  $t$ -test (\* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ )

#### 4 Using PID to measure stimulus-related synergy and redundancy in auditory cortex during correct and incorrect behaviour

However, the above results leave the question of how synergy and redundancy separately change between correct and incorrect trials unaddressed. Thus, it is not clear how synergy and redundancy correlate with the accuracy of behavioural decisions.

In fact, it has been shown that co-information conflates two non-negative pieces of information which properly and separately quantify synergy and redundancy [30]. Indeed, there could be cases in which co-information is low, but synergy and redundancy are both high and cancel out due to their opposing signs [43]. In simple terms, redundancy (the area in red in the Venn diagram in Fig. 2B) quantifies the amount of information that both neurons carry independently about the stimulus, whilst synergy (the area in green in the Venn diagram in Fig. 2B) is the amount of information that can be accessed when observing both neuronal responses simultaneously and is not carried individually by any of the two neurons. Thus, the previously reported results could arise in distinct scenarios: redundancy is higher in correct rather than in incorrect trials, synergy is lower in correct than in incorrect trials, or a combination of the two.

To determine the specific contributions of synergy and redundancy to the total joint information, we used the formalism of PID [30]. PID allows breaking down the joint mutual information that two or more source variables carry about a target variable into non-negative and interpretable pieces of information (termed *information atoms*) which quantify how information about the target variable is distributed amongst source variables. In the case of a system with two source variables and one target variable, PID breaks down the joint mutual information encoded by the two sources about the target (See Eq. (2)) into four non-negative information atoms [30, 44]:

$$SI(S; R_i, R_j) = Red(S : R_i, R_j) + Syn(S : R_i, R_j) + Un_i(S : R_i \setminus R_j) + Un_j(S : R_j \setminus R_i), \quad (4)$$

where  $Red(S : R_i, R_j)$  is the redundant information (red area in the Venn diagram in Fig. 2B) which is present in both neuron  $R_i$  and neuron  $R_j$ ,  $Syn(S : R_i, R_j)$  is the synergistic information (green area in Fig. 2B) carried only by the joint response of the two neurons, whilst  $Un_i(S : R_i \setminus R_j)$  and  $Un_j(S : R_j \setminus R_i)$  stand for the two unique information components (grey and white areas in Fig. 2B, respectively) carried by one source variable but not by the other. Importantly, the four information atoms appearing in the right-hand side of Eq. (4) are not independent, so that determining the value of one atom is sufficient to compute all the others, as the other three can be

computed as linear combinations of Shannon information-theoretic quantities and the determined atom [44].

An important insight arising from the PID is that  $CoInfo(S; R_i; R_j)$  given in Eq. (3) is the difference between the two distinct information atoms that express synergy and redundancy, respectively:

$$CoInfo(S; R_i; R_j) = Syn(S : R_i, R_j) - Red(S : R_i, R_j). \quad (5)$$

To compute  $Red(S : R_i, R_j)$  and  $Syn(S : R_i, R_j)$  we used the definition provided by [44]. Given a trivariate probability distribution  $P(S, R_i, R_j)$ , Bertschinger et al. defined the unique information atom as follows [44]:

$$Un_i(S : R_i, R_j) = \min_{Q \in \Delta_P} I_Q(S : R_i \setminus R_j), \quad (6)$$

which defines a constrained convex optimization problem in the space  $\Delta_P$  of trivariate probability distributions  $Q(S, R_i, R_j)$  with fixed marginals  $Q(S, R_i) = P(S, R_i)$  and  $Q(S, R_j) = P(S, R_j)$ .

To numerically solve this optimization problem, we used the *BROJA\_2PID* python package [45]. In this way, we computed the synergistic and the redundant information that pairs of neurons carried about the stimulus.

Following [34], we labelled the neuronal pairs as GC-connected if they shared at least one GC link and as GC-unconnected otherwise. We performed the stimulus-related PID analysis in the two separate groups of GC-connected and GC-unconnected pairs of neurons in correct and incorrect trials.

We first performed the PID analysis in correct trials for both the GC-connected and GC-unconnected pairs of neurons (Fig. 3A). To obtain a fair comparison between results in correct and incorrect trials, we performed these analyses over a randomly selected subsample of correct trials with the same sample size as the incorrect trials (results are presented as average over 100

random subsamples). We found that the joint stimulus information had comparable values ( $0.386 \pm 0.002$  bits vs  $0.402 \pm 0.012$  bits for GC-unconnected vs GC-connected pairs respectively; hereafter, all results in this section are reported as mean  $\pm$  SEM over all pairs of neurons) in both populations. However, GC-connected pairs had higher levels of redundancy ( $0.121 \pm 0.006$  bits) compared to the GC-unconnected ones ( $0.105 \pm 0.001$  bits), whilst they had similar amounts of synergy for GC-unconnected ( $0.097 \pm 0.001$  bits) and GC-connected pairs ( $0.093 \pm 0.001$  bits) respectively (Fig. 3A). Confirming the previously reported results [34], the difference

between synergy and redundancy, i.e., the co-information (Eq. (5)), showed a prevalence of redundant information in both populations, but the GC-connected pairs were more redundant ( $-0.027 \pm 0.006$  bits) than GC-unconnected pairs ( $-0.007 \pm 0.001$  bits).

We next quantified the fraction of redundancy and synergy by normalizing each term with respect to the total joint mutual information. This is useful to discount any possible effect of differences in information levels between correct and error trials. We found that GC-connected pairs had proportionally more redundancy and less synergy ( $\text{Red}=0.292 \pm 0.010$ ,  $\text{Syn}=0.239 \pm 0.007$ ), compared to GC-unconnected ones ( $\text{Red}=0.262 \pm 0.001$ ,  $\text{Syn}=0.261 \pm 0.001$ ) (Fig. 3A). Moreover, GC-connected pairs had much more predominant redundancy ( $-0.053 \pm 0.013$ ) than GC-unconnected pairs ( $-0.001 \pm 0.002$ ).

In sum, our results suggest that GC-connected pairs of neurons have more redundant than synergistic functional connections.

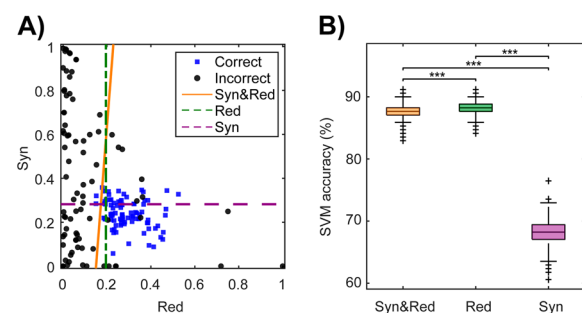
Next, we investigated whether higher amounts of redundancy and lower amounts of synergy could be beneficial for task performance and behavioural accuracy. We computed the PID in incorrect trials (Fig. 3B). The joint stimulus information in incorrect trials ( $0.130 \pm 0.002$  bits,  $0.123 \pm 0.014$  bits for GC-unconnected and GC-connected pairs respectively) was only  $\sim 30\%$  of what it was in correct trials. Redundancy in incorrect trials had a value of  $0.010 \pm 0.001$  bits,  $0.012 \pm 0.004$  bits for GC-unconnected and GC-connected pairs respectively, which is proportionally 10 times smaller than that of correct trials. Synergy dropped to  $0.063 \pm 0.002$  bits and  $0.053 \pm 0.007$  bits for GC-unconnected and GC-connected pairs respectively, proportionally only half of that in correct trials. Co-information showed positive values, i.e., more synergy than redundancy, in both GC-unconnected ( $0.053 \pm 0.002$  bits) and GC-connected pairs ( $0.040 \pm 0.007$  bits). Normalized redundancy constituted approximately 10% of the total information, whereas normalized synergy amounted to  $\sim 45\%$  (Fig. 3B). We did not find significant differences in the normalized co-information between GC-unconnected and GC-connected pairs on incorrect trials ( $0.382 \pm 0.008$  vs  $0.304 \pm 0.044$ ). Our results suggest that only the redundant FC associated with GC links is beneficial to correct sensory discrimination.

## 5 Predicting correct vs incorrect perceptual discriminations based on redundancy and synergy of functionally connected neurons

Given that GC-connected pairs of neurons exhibit higher values of normalized redundancy and lower values of normalized synergy during correct decisions compared

to incorrect ones, we sought to determine whether redundancy or synergy is more predictive of the correctness of perceptual discriminations. We focused this analysis on normalized redundancy and synergy values to control for potential confounding effect of differences in joint information values between correct and incorrect trials.

To visualize this dependency in an intuitive way, in Fig. 4A we present a scatterplot (on the  $n=85$  pairs of GC-connected neurons that were used in previous analyses) of how normalized synergy and redundancy values are distributed across GC-connected pairs for correct and incorrect decisions. Visual inspection of this plot suggests that normalized redundancy values have a strong discrimination power, with high values of normalized redundancy predicting correct choices and low values of normalized redundancy predicting incorrect choices. To support this intuition with quantitative analyses, we used a soft-margin Support Vector Machine (SVM) with a linear kernel to discriminate between correct and incorrect behavioural decisions from the normalized synergy and redundancy values of GC-connected pairs. Specifically, we used the MATLAB function *fitcsvm* with default arguments, in particular with regularization parameter  $C=1$  and optimization with the Sequential Minimal Optimization algorithm, and we used a leave-one-out cross-validation procedure on the  $n=85$  pairs of GC-connected neurons. We first classified correct



**Fig. 4** Decoding behavioural decisions through normalized synergy and redundancy in GC-connected pairs. **A** We used normalized synergy and redundancy to classify behavioural decisions as correct (blue squares) or incorrect (black dots) perceptual discriminations. The solid orange line is the decision boundary of an SVM trained using both synergy and redundancy values as predictive features. Green and purple dashed lines are the decision boundaries of SVM trained using only redundancy or synergy, respectively, as a predictive feature. Correct decisions are characterized by clustered values in both synergy and redundancy, whilst incorrect decisions show lower redundancy values and a uniform distribution of synergy. **B** SVM cross-validated classification accuracy from 10,000 bootstrap samples. Colours of the distributions match the ones of decision boundaries in **A**. Statistical comparisons across different models were made using the Wilcoxon rank-sum test (\*\*\*)  $p < 0.001$

vs incorrect behaviour when the SVM used both normalized redundancy and synergy and found that, when using both features, it yielded a high classification accuracy of  $87.60 \pm 1.20\%$  (hereafter, in this section values are reported as mean  $\pm$  SD across 10,000 bootstrap samples) for correct vs incorrect decisions (Fig. 4B). The SVM weight corresponding to the normalized redundancy had a higher absolute magnitude than the one corresponding to the normalized synergy (magnitude of the redundancy SVM weight:  $5.79 \pm 0.05$ ; magnitude of the synergy SVM weight:  $-0.65 \pm 0.04$ ). These results indicate that proportion of redundancy has a major predictive power for the correctness of behavioural decisions.

To further assess the roles of normalized synergy and redundancy values in predicting correct decisions, we used the SVM to classify error vs correct behaviour using each feature separately. The classifier using exclusively on normalized redundancy discriminated correct vs incorrect decisions with a high accuracy ( $88.28 \pm 0.97\%$ ), even slightly higher than the classification accuracy achieved considering both normalized synergy and redundancy (Fig. 4B). In contrast, the SVM classifier relying solely on normalized synergy achieved a lower accuracy ( $67.98 \pm 1.77\%$ ) than both the classifiers relying solely on synergy or on both synergy and redundancy (Fig. 4B). This shows that once redundancy is known, synergy does not add predictive power about the correctness of perceptual discriminations.

## 6 Discussion

In this study, we teased apart the relationship between FC and stimulus-related synergy and redundancy with single-neuron resolution in the mouse auditory cortex during a perceptual discrimination task. We deliberately considered one specific, widely used type of directed FC measure, Granger Causality. GC is a directed measure, and as such it can disambiguate between stronger information transfer in one direction than the opposite direction. It is a data-robust linear version of the corresponding information theoretic quantity, Transfer Entropy (TE) [46]. Whilst TE has the advantage of possibly capturing non-linear information transfer and it can be also framed in the context of PID [28, 47], the data-robustness of GC allows its easier application in multivariate settings to condition away the effects of other neurons [26]. Importantly for the present study, unlike other measures such as the Pearson correlation between the activity of two neurons, GC can in principle be related to both redundancy and synergy.

Our findings revealed that Granger FC between A1 L2/3 neurons was accompanied by proportionally higher levels of redundancy and lower levels of synergy compared to pairs of neurons that were not linked with a

Granger FC. These results suggest that FC creates prevalent redundancy of sensory information across neurons.

Previous work has established that the sensory information encoded by neuronal populations greatly decreases when animals make incorrect perceptual decisions, compared to when animals make correct decisions [17, 48–52]. However, less is still known about how the interactions between neurons in a population code, and the patterns of synergy and redundancy that may be created by these interactions, promote correct decisions [4]. Here, we made progress in this direction by studying not only how information levels change between correct and incorrect trials but also studying patterns of synergy and redundancy. Our results suggest that both synergy and redundancy coexist across the population, both when mice make correct or incorrect perceptual discriminations. However, we found that the levels of redundancy were much higher (both in absolute terms and in proportion to the total information available in neuron pairs) in both populations when mice made correct behavioural choices compared to incorrect ones, whereas synergy values were higher in absolute terms but lower in relative terms during correct compared to incorrect behavioural choices. Moreover, the proportion of redundancy more reliably predicted perceptual discriminations, whilst the proportion of synergy had a much lower predictive power per se, and did not add predictive power once the proportion of redundancy was known.

Overall, the above results suggest that redundancy is highly beneficial for correct sensory judgements. The advantages of redundancy for perceptual discrimination could arise from multiple contributions. One well-documented advantage regards the integration of information across sites [53]. Another one could result in advantages in terms of information transmission and readout. Indeed, whilst redundancy limits the amount of encoded information [54], it has benefits in terms of improving the propagation of information between pre- and post-synaptic neurons [4, 17]. Together with those reported in previous studies [4, 17, 34], our results suggest that the optimal trade-off between the advantages and disadvantages of redundancy results in an overall advantage of having some degree of redundancy to secure reliable downstream information transmission.

Our findings confirm previous reports of significant synergy between the activity of neurons or networks [14, 15, 33]. Our finding that synergy is higher in absolute terms during correct behaviour suggest that synergy may promote correct decisions by elevating the information levels in correct trials. However, our observations of a decreased proportion of synergy during correct perceptual discrimination suggests that the potential advantage of synergy in terms of higher levels of sensory

information encoding may not entirely translate into an advantage for sensory discrimination. One possibility is that the interactions leading to synergistic information may be more difficult to be read out by downstream computations, as they would require more sophisticated decoders that may be beyond the capabilities of some downstream neural circuits. However, given that presence of synergy has been well-documented, another possibility, to be explored in future studies, is that synergy may not be needed for the simple perceptual tasks we consider and for neurons in sensory areas, but that it could become more important for more complex behaviours or for neurons in higher level areas [14].

Together, these results establish the major importance of redundancy amongst neurons in sensory cortices for correct sensory discriminations which may be due to the beneficial effects that redundancy has on downstream information transmission. At the same time, our results suggest also a smaller yet useful contribution of synergy to correct perceptual discriminations, by enhancing information levels during correct behaviour.

Another important question regards how synergistic and redundant FC relate to structural connectivity [14]. Robust and meaningful relationships have been established between redundant FC measured during the resting state and structural connectivity at the level of whole-brain measures that lack cellular resolution [14, 21]. However, it remains to be understood how this anatomical substrate is complemented by stimulus-dependent changes in neural dynamics. The same structural connectivity can give rise to different patterns of functional connectivity depending for example on the state of each node. For example, depending on the degree of excitability of a given node, the functional interactions between areas can be larger or smaller even if the anatomical connections between them do not change. As a result, the relationship between functional and structural connectivity is complex [55] and changes in state of individual nodes or on the stimulus information present in the inputs to some of the considered nodes can modulate both redundancy and synergy between anatomically connected nodes. Detailed studies of realistic neural network models, as well as careful experiments that manipulate the activity of individual nodes [56], will be a key to progress in addressing these questions.

From a theoretical perspective, previous studies that investigated synergy and redundancy between neurons or networks employed a measure of co-information which conflates synergy with redundancy, measuring only their net effect [32, 34]. Our work advances the state-of-the-art by providing a more refined measure that delineates redundancy from synergy and enables separate quantification of their relationship with both FC and

the accuracy of behaviour. With respect to other studies considering redundancy and synergy, but not relating it to information content about variables of cognitive interest [14], we made progress by measuring redundancy and synergy of information about variables, such as sensory stimuli, which have a well-defined meaning and role in terms of perceptual functions. We hope that our work will contribute to creating a neuroinformatics framework that can help researchers to study the patterns of synergy and redundancy about external stimuli and pinpoint their contribution to behaviour and functions. This progress will need to include mathematical advances in the understanding of the differences and complementarity between different possible PID formalisms. For example, the formalism we used here [44] breaks down information into non-negative parts, as in the PID original formulation [30, 47]. However other work is exploring the advantages of alternative ways to decompose information, including decompositions into terms that do not need to be non-negative [57–60]. It would also be important to understand the relationship between the PID-based formalisms and previously derived information-theoretic formalisms that quantify how the information in a population of neurons depends on the correlations of the activity of different neurons [32, 61–63]. These previous studies established important rules for how correlations between neurons can enhance or decrease information and change co-information values (for example, correlations can increase information and thus create synergy, when their strength is modulated by the stimulus). Connecting these formalisms will aid the understanding of how synergy and redundancy may arise in terms of basic properties of neural activity or of circuit mechanisms.

In conclusion, our study provides a framework to measure the behavioural relevance of synergy and redundancy even with cellular resolution. The results obtained analysing the activity of auditory cortex with this framework suggest that correct behaviour is associated with a predominant presence of redundant information in functionally connected neural networks. Further research is needed to better understand the contributions of synergy and redundancy in different contexts.

#### Abbreviations

|        |                                       |
|--------|---------------------------------------|
| FC     | Functional connectivity               |
| fMRI   | Functional Magnetic Resonance Imaging |
| EEG    | Electroencephalogram                  |
| GC     | Granger causality                     |
| PID    | Partial Information Decomposition     |
| A1     | Primary auditory cortex               |
| L2/3   | Layer 2/3                             |
| SEM    | Standard Error of the Mean            |
| SI     | Stimulus information                  |
| Colnfo | Co-information                        |
| Red    | Redundancy                            |

|     |                        |
|-----|------------------------|
| Syn | Synergy                |
| Un  | Unique information     |
| SVM | Support Vector Machine |
| SD  | Standard deviation     |
| TE  | Transfer entropy       |

### Acknowledgements

We are most grateful to the organizers and participants of the 16th International Conference on Brain Informatics (BI 2023) for their feedback on this work.

### Author contributions

LK and SP conceived the study. SP supervised the study. LK performed information theoretic analyses. NAF performed experiments. NAF, MC, SM, LK, BB, POK and SP provided methods. LK, MC, and SP wrote the paper with input from all authors.

### Funding

Open Access funding enabled and organized by Projekt DEAL. This research was supported by National Institutes of Health (NIH) Brain Initiative grants R01 NS109961 (SP) and R01 NS108410 (SP), by an (NIH) Brain Initiative grant U19 NS107464 (BB, POK and SP), and by National Science Foundation grants ECCS1807216 (BB) and ECCS2032649 (BB).

### Availability of data and materials

The code used to compute information measures is taken from [40] and can be downloaded at <https://doi.org/10.5281/zenodo.850362>. The experimental data were shared in a previous publication [34] and can be downloaded from the Digital Repository at the University of Maryland at <https://drum.lib.umd.edu/items/30d43732-7149-4726-a860-0ae3d210b2ae>. Any additional information required to reanalyse the data reported in this paper is available from the corresponding authors upon reasonable request.

### Declarations

### Competing interests

The authors declare that they have no competing interests.

### Author details

<sup>1</sup>Istituto Italiano Di Tecnologia, 38068 Rovereto, Italy. <sup>2</sup>Department of Excellence for Neural Information Processing, Center for Molecular Neurobiology (ZMNH), University Medical Center Hamburg-Eppendorf (UKE), Falkenried 94, 20251 Hamburg, Germany. <sup>3</sup>Department of Neurophysiology and Pathophysiology, University Medical Center Hamburg-Eppendorf (UKE), 20246 Hamburg, Germany. <sup>4</sup>Department of Pharmacy and Biotechnology, University of Bologna, 40126 Bologna, Italy. <sup>5</sup>Department of Biology and Brain and Behavior Institute, University of Maryland, College Park, MD 20742, USA. <sup>6</sup>Department of Electrical and Computer Engineering and Institute for Systems Research, University of Maryland, College Park, MD 20742, USA. <sup>7</sup>Department of Biomedical Engineering and Kavli Neuroscience Discovery Institute, Johns Hopkins University, Baltimore, MD 21205, USA.

Received: 6 October 2023 Accepted: 2 November 2023

Published online: 05 December 2023

### References

- Biswal B, Yetkin FZ, Haughton VM, Hyde JS (1995) Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magn Reson Med* 34(4):537–541. <https://doi.org/10.1002/mrm.1910340409>
- Greicius MD, Krasnow B, Reiss AL, Menon V (2003) Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. *Proc Natl Acad Sci USA* 100(1):253–258. <https://doi.org/10.1073/pnas.0135058100>
- Fox MD, Snyder AZ, Vincent JL, Corbetta M, Van Essen DC, Raichle ME (2005) The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proc Natl Acad Sci USA* 102(27):9673–9678. <https://doi.org/10.1073/pnas.0504136102>
- Panzeri S, Moroni M, Safaai H, Harvey CD (2022) The structures and functions of correlations in neural population codes. *Nat Rev Neurosci* 23(9):551–567. <https://doi.org/10.1038/s41583-022-00606-4>
- Engel AK, Gerloff C, Hülsmann CC, Nolte G (2013) Intrinsic coupling modes: multiscale interactions in ongoing brain activity. *Neuron* 80(4):867–886. <https://doi.org/10.1016/j.neuron.2013.09.038>
- Hutchison RM, Womelsdorf T, Allen EA, Bandettini PA, Calhoun VD, Corbetta M, Della Penna S, Duyn JH, Glover GH, Gonzalez-Castillo J, Handwerker DA, Keilholz S, Kiviniemi V, Leopold DA, de Pasquale F, Sporns O, Walter M, Chang C (2013) Dynamic functional connectivity: promise, issues, and interpretations. *Neuroimage* 80:360–378. <https://doi.org/10.1016/j.neuroimage.2013.05.079>
- Vincent JL, Patel GH, Fox MD, Snyder AZ, Baker JT, Van Essen DC, Zempel JM, Snyder LH, Corbetta M, Raichle ME (2007) Intrinsic functional architecture in the anaesthetized monkey brain. *Nature* 447(7140):83–86. <https://doi.org/10.1038/nature05758>
- Gozzi A, Schwarz AJ (2016) Large-scale functional connectivity networks in the rodent brain. *Neuroimage* 127:496–509. <https://doi.org/10.1016/j.neuroimage.2015.12.017>
- Fox MD, Greicius M (2010) Clinical applications of resting state functional connectivity. *Front Syst Neurosci* 4:19. <https://doi.org/10.3389/fnsys.2010.00019>
- Bertero A, Liska A, Pagani M, Parolisi R, Masferrer ME, Gritti M, Pedrazzoli M, Galbusera A, Sarica A, Cerasa A, Buffelli M, Tonini R, Buffo A, Gross C, Pasqualetti M, Gozzi A (2018) Autism-associated 16p112 microdeletion impairs prefrontal functional connectivity in mouse and human. *Brain* 141(7):2055–2065. <https://doi.org/10.1093/brain/awy111>
- Mediano PAM, Rosas FE, Luppi AI, Jensen HJ, Seth AK, Barrett AB, Carhart-Harris RL, Bor D (2022) Greater than the parts: a review of the information decomposition approach to causal emergence. *Phil Trans Roy Soc A* 380(2227):20210246. <https://doi.org/10.1098/rsta.2021.0246>
- Newman EL, Varley TF, Parakkattu VK, Sherrill SP, Beggs JM (2022) Revealing the dynamics of neural information processing with multivariate information decomposition. *Entropy* 24(7):930. <https://doi.org/10.3390/e24070930>
- Varley TF, Pope M, Faskowitz J, Sporns O (2023) Multivariate information theory uncovers synergistic subsystems of the human cerebral cortex. *Commun Biol* 6:451. <https://doi.org/10.1038/s42003-023-04843-w>
- Luppi AI, Mediano PAM, Rosas FE, Holland N, Fryer TD, O'Brien JT, Rowe JB, Menon DK, Bor D, Stamatakis EA (2022) A synergistic core for human brain evolution and cognition. *Nat Neurosci* 25(6):771–782. <https://doi.org/10.1038/s41593-022-01070-0>
- Varley TF, Sporns O, Schaffelhofer S, Scherberger H, Dann B (2023) Information-processing dynamics in neural networks of macaque cerebral cortex reflect cognitive state and behavior. *Proc Natl Acad Sci USA* 120(2):e2207677120. <https://doi.org/10.1073/pnas.2207677120>
- Sporns O (2022) The complex brain: connectivity, dynamics, information. *Trends Cogn Sci* 26(12):1066–1067. <https://doi.org/10.1016/j.tics.2022.08.002>
- Valente M, Pica G, Bondanelli G, Moroni M, Runyan CA, Morcos AS, Harvey CD, Panzeri S (2021) Correlations enhance the behavioral readout of neural population activity in association cortex. *Nat Neurosci* 24(7):975–986. <https://doi.org/10.1038/s41593-021-00845-1>
- Gatica M, Cofre R, Mediano PAM, Rosas FE, Orio P, Diez I, Swinnen SP, Cortes JM (2021) High-Order Interdependencies in the Aging Brain. *Brain Connect* 11(9):734–744. <https://doi.org/10.1089/brain.2020.0982>
- van den Heuvel MP, Hulshoff Pol HE (2010) Exploring the brain network: a review on resting-state fMRI functional connectivity. *Eur Neuropsychopharmacol* 20(8):519–534. <https://doi.org/10.1016/j.euroneuro.2010.03.008>
- Deco G, Ponce-Alvarez A, Mantini D, Romani GL, Hagmann P, Corbetta M (2013) Resting-state functional connectivity emerges from structurally and dynamically shaped slow linear fluctuations. *J Neurosci* 33(27):11239–11252. <https://doi.org/10.1523/JNEUROSCI.1091-13.2013>
- Honey CJ, Sporns O, Cammoun L, Gigandet X, Thiran JP, Meuli R, Hagmann P (2009) Predicting human resting-state functional connectivity from structural connectivity. *Proc Natl Acad Sci USA* 106(6):2035–2040. <https://doi.org/10.1073/pnas.0811168106>
- Lachaux JP, Rodriguez E, Martinerie J, Varela FJ (1999) Measuring phase synchrony in brain signals. *Hum Brain Mapp* 8(4):194–208. [https://doi.org/10.1002/\(sici\)1097-0193\(1999\)8:4%3c194::aid-hbm4%3e3.0.co;2-c](https://doi.org/10.1002/(sici)1097-0193(1999)8:4%3c194::aid-hbm4%3e3.0.co;2-c)

23. Nolte G, Bai O, Wheaton L, Mari Z, Vorbach S, Hallett M (2004) Identifying true brain interaction from EEG data using the imaginary part of coherence. *Clin Neurophysiol* 115(10):2292–2307. <https://doi.org/10.1016/j.clinph.2004.04.029>
24. Sherrill SP, Timme NM, Beggs JM, Newman EL (2021) Partial information decomposition reveals that synergistic neural integration is greater downstream of recurrent information flow in organotypic cortical cultures. *PLoS Comput Biol* 17(7):e1009196. <https://doi.org/10.1371/journal.pcbi.1009196>
25. Seth AK, Barrett AB, Barnett L (2015) Granger causality analysis in neuroscience and neuroimaging. *J Neurosci* 35(8):3293–3297. <https://doi.org/10.1523/JNEUROSCI.4399-14.2015>
26. Sheikhattar A, Miran S, Liu J, Fritz JB, Shamma SA, Kanold PO, Babadi B (2018) Extracting neuronal functional network dynamics via adaptive Granger causality analysis. *Proc Natl Acad Sci USA* 115(17):E3869–E3878. <https://doi.org/10.1073/pnas.1718154115>
27. Schreiber T (2000) Measuring information transfer. *Phys Rev Lett* 85(2):461–464. <https://doi.org/10.1103/PhysRevLett.85.461>
28. Celotto M, Birn J, Tlaie A, De Feo V, Lemke S, Chicharro D, Nili H, Bieler M, Hanganu-Opatz IL, Donner TH, Brovelli A, Panzeri S (2023) An information-theoretic quantification of the content of communication between brain regions. *Adv Neural Inf Process Syst (NeurIPS)* 37 (in press). <https://neurips.cc/virtual/2023/poster/70605>
29. Besserve M, Lowe SC, Logothetis NK, Schölkopf B, Panzeri S (2015) Shifts of gamma phase across primary visual cortical sites reflect dynamic stimulus-modulated information transfer. *PLOS Biol* 13(9):e1002257. <https://doi.org/10.1371/journal.pbio.1002257>
30. Williams PL, Beer RD (2010) Nonnegative decomposition of multivariate information. <https://doi.org/10.48550/arXiv.1004.2515>
31. Wibral M, Priesemann V, Kay JW, Lizier JT, Phillips WA (2017) Partial information decomposition as a unified approach to the specification of neural goal functions. *Brain Cogn* 112:25–38. <https://doi.org/10.1016/j.bandc.2015.09.004>
32. Schneidman E, Bialek W, Berry MJ (2003) Synergy, redundancy, and independence in population codes. *J Neurosci* 23(37):11539–11553. <https://doi.org/10.1523/JNEUROSCI.23-37-11539.2003>
33. Nigam S, Pojoga S, Dragoi V (2019) Synergistic coding of visual information in columnar networks. *Neuron* 104(2):402–411. <https://doi.org/10.1016/j.neuron.2019.07.006>
34. Francis NA, Mukherjee S, Koçillari L, Panzeri S, Babadi B, Kanold PO (2022) Sequential transmission of task-relevant information in cortical neuronal networks. *Cell Rep* 39(9):110878. <https://doi.org/10.1016/j.celrep.2022.110878>
35. Koçillari L, Celotto M, Francis NA, Mukherjee S, Babadi B, Kanold PO, Panzeri S (2023) Measuring Stimulus-Related Redundant and Synergistic Functional Connectivity with Single Cell Resolution in Auditory Cortex. In: Liu F, Zhang Y, Kuai H, Stephen EP and Wang H (eds) *Brain Informatics. BI 2023. Lecture Notes in Computer Science*. 13974. Springer, Cham, pp. 45–56 [https://doi.org/10.1007/978-3-031-43075-6\\_5](https://doi.org/10.1007/978-3-031-43075-6_5)
36. Shannon CE (1948) A mathematical theory of communication. *Bell Syst Tech J* 27(3):379–423. <https://doi.org/10.1002/j.1538-7305.1948.tb01338.x>
37. Quiñero R, Panzeri S (2009) Extracting information from neuronal populations: information theory and decoding approaches. *Nat Rev Neurosci* 10(3):173–185. <https://doi.org/10.1038/nrn2578>
38. Magri C, Whittingstall K, Singh V, Logothetis NK, Panzeri S (2009) A toolbox for the fast information analysis of multiple-site LFP, EEG and spike train recordings. *BMC Neurosci* 10:81. <https://doi.org/10.1186/1471-2202-10-81>
39. Panzeri S, Senatore R, Montemurro MA, Petersen RS (2007) Correcting for the sampling bias problem in spike train information measures. *J Neurophysiol* 98(3):1064–1072. <https://doi.org/10.1152/jn.00559.2007>
40. Pica G, Piasini E, Safaai H, Runyan CA, Diamond ME, Fellin T, Kayser C, Harvey CD, Panzeri S (2017) Quantifying how much sensory information in a neural code is relevant for behavior. In: Guyon I, Luxburg UV, Bengio S, Wallach H, Fergus R, Vishwanathan S and Garnett R (eds) *Adv Neural Inf Process Syst (NeurIPS)*. 30. Curran Associates, Inc, pp. 3689–3699.
41. Panzeri S, Harvey CD, Piasini E, Latham PE, Fellin T (2017) Cracking the neural code for sensory perception by combining statistics, intervention, and behavior. *Neuron* 93(3):491–507. <https://doi.org/10.1016/j.neuron.2016.12.036>
42. McGill WJ (1954) Multivariate information transmission. *Psychometrika* 19(2):97–116. <https://doi.org/10.1007/BF02289159>
43. Griffith V, Koch C (2014) Quantifying Synergistic Mutual Information. In: Prokopenko M (eds) *Guided self-organization: inception, emergence, complexity and computation*. 9. Springer, Berlin, Heidelberg, pp. 159–190 [https://doi.org/10.1007/978-3-642-53734-9\\_6](https://doi.org/10.1007/978-3-642-53734-9_6)
44. Bertschinger N, Rauh J, Olbrich E, Jost J, Ay N (2014) Quantifying unique information. *Entropy* 16(4):2161–2183. <https://doi.org/10.3390/e16042161>
45. Makkeh A, Theis DO, Vicente R (2018) BROJA-2PID: a robust estimator for bivariate partial information decomposition. *Entropy* 20(4):271. <https://doi.org/10.3390/e20040271>
46. Barnett L, Barrett AB, Seth AK (2009) Granger causality and transfer entropy are equivalent for gaussian variables. *Phys Rev Lett* 103(23):238701. <https://doi.org/10.1103/PhysRevLett.103.238701>
47. Williams PL, Beer RD (2011) Generalized Measures of Information Transfer. <https://doi.org/10.48550/arXiv.1102.1507>
48. Luna R, Hernández A, Brody CD, Romo R (2005) Neural codes for perceptual discrimination in primary somatosensory cortex. *Nat Neurosci* 8(9):1210–1219. <https://doi.org/10.1038/nn1513>
49. Rigotti M, Barak O, Warden MR, Wang X-J, Daw ND, Miller EK, Fusi S (2013) The importance of mixed selectivity in complex cognitive tasks. *Nature* 497:585–590. <https://doi.org/10.1038/nature12160>
50. Zuo Y, Safaai H, Notaro G, Mazzoni A, Panzeri S, Diamond ME (2015) Complementary contributions of spike timing and spike rate to perceptual decisions in rat S1 and S2 cortex. *Curr Biol* 25(3):357–363. <https://doi.org/10.1016/j.cub.2014.11.065>
51. Runyan CA, Piasini E, Panzeri S, Harvey CD (2017) Distinct timescales of population coding across cortex. *Nature* 548(7665):92–96. <https://doi.org/10.1038/nature23020>
52. Kira S, Safaai H, Morcos AS, Panzeri S, Harvey CD (2023) A distributed and efficient population code of mixed selectivity neurons for flexible navigation decisions. *Nat Commun* 14(1):2121. <https://doi.org/10.1038/s41467-023-37804-2>
53. Tononi G, Sporns O, Edelman GM (1994) A measure for brain complexity: relating functional segregation and integration in the nervous system. *Proc Natl Acad Sci USA* 91(11):5033–5037. <https://doi.org/10.1073/pnas.91.11.5033>
54. Averbeck BB, Latham PE, Pouget A (2006) Neural correlations, population coding and computation. *Nat Rev Neurosci* 7(5):358–366. <https://doi.org/10.1038/nrn1888>
55. Celotto M, Lemke S, Panzeri S (2022) Inferring the temporal evolution of synaptic weights from dynamic functional connectivity. *Brain Inf* 9:28. <https://doi.org/10.1186/s40708-022-00178-0>
56. Rocchi F, Canella C, Noei S, Gutierrez-Barragan D, Coletta L, Galbusera A, Stuefer A, Vassanelli S, Pasqualetti M, Iurilli G, Panzeri S, Gozzi A (2022) Increased fMRI connectivity upon chemo-genetic inhibition of the mouse prefrontal cortex. *Nat Commun* 13(1):1056. <https://doi.org/10.1038/s41467-022-28591-3>
57. Makkeh A, Gutknecht AJ, Wibral M (2021) Introducing a differentiable measure of pointwise shared information. *Phys Rev E* 103(3):032149. <https://doi.org/10.1103/PhysRevE.103.032149>
58. Finn C, Lizier JT (2018) Pointwise partial information decomposition using the specificity and ambiguity lattices. *Entropy* 20(4):297. <https://doi.org/10.3390/e20040297>
59. Kolchinsky A (2022) A novel approach to the partial information decomposition. *Entropy* 24(3):403. <https://doi.org/10.3390/e24030403>
60. Ince RAA (2017) Measuring multivariate redundant information with pointwise common change in surprisal. *Entropy* 19(7):318. <https://doi.org/10.3390/e19070318>
61. Panzeri S, Schultz SR, Treves A, Rolls ET (1999) Correlations and the encoding of information in the nervous system. *Proc Biol Sci* 266(1423):1001–1012. <https://doi.org/10.1098/rspb.1999.0736>
62. Pola G, Thiele A, Hoffmann KP, Panzeri S (2003) An exact method to quantify the information transmitted by different mechanisms of correlational coding. *Network* 14(1):35–60. <https://doi.org/10.1088/0954-898x/14/1/303>
63. Latham PE, Nirenberg S (2005) Synergy, redundancy, and independence in population codes, revisited. *J Neurosci* 25(21):5195–5206. <https://doi.org/10.1523/JNEUROSCI.5319-04.2005>

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.