Modulation of Brain Connectivity by Memory Load in a Working Memory Network

Pouya Bashivan*
Mohammed Yeasin
Department of Electrical and Computer Engineering
The University of Memphis
Memphis, TN, USA
{pbshivan;myeasin}@memphis.edu

Gavin M. Bidelman
Institute for Intelligent Systems and
School of Communication Sciences & Disorders
The University of Memphis
Memphis, TN, USA
gmbdlman@memphis.edu

***Abstract*— Cognition is the product of activation of billions of neurons and their timely interactions. While the activity of individual neurons is essential for proper functioning of the brain, the communication among them is arguably more vital. Previous studies of brain connectivity have largely focused on investigating causality across the brain in order to reveal the existing communication channels that form its internal networks. However, little is known about how these neuronal pathways respond to task demands with varying degrees of complexity. Towards understanding the pathways of information flow, we investigated the effect of memory load on network connectivity of brain. Independent component analysis (ICA) was used to identify brain areas, active during a working memory task, whose activations co-varied with memory load. An information theoretic metric called *transfer entropy* was adopted to examine the directed links across these areas. Empirical results suggest that the information flow rate across a primary working memory network is modulated by memory load. Furthermore, it was observed that the information flow is affected in pathways with opposite direction during *encoding* and *maintenance* stages of working memory operation.**

Keywords—effective connectivity, EEG, independent component analysis, transfer entropy, working memory

I. INTRODUCTION

Information processing within the human brain can be envisioned as a network of sub-processes. The structure of this network has been shown to possess features of complex networks such as small-world topology, highly connected hubs and modularity [1]. Phase-locked electromagnetic oscillations have been proposed as a possible communication method across the activated neuronal groups [2]. While localized activity of the brain has been studied for decades, brain connectivity, which studies the possible relationship between different areas of brain, is a relatively recent advance. The human connectome has been investigated using a range of diverse modalities including structural and functional MRI (fMRI), diffusion tensor imaging (DTI), magnetoencephalography (MEG) and electroencephalography (EEG). In addition, different mathematical algorithms including Granger causality (GC), dynamic causal modeling (DCM), transfer entropy (TE), and phase synchronization have been utilized to study the functional/effective connectivity across the brain. A thorough review of the different modalities and algorithms can be found in [3] and [1].

While widely used methods like GC and DCM depend on specific model structures or *a priori* knowledge about the system, transfer entropy is a non-parametric method and does not require any *a priori* information about the underlying process. These characteristics make TE particularly suitable for exploratory analyses [4]. Furthermore, TE can detect nonlinear interactions between the signals and behaves robustly against linear cross-talk which is a common issue in EEG/MEG signals due to their volume conducted nature [4]. The TE has previously been used to detect task-related information flow and to predict task dependent interdependencies using MEG and EEG data [5], [6].

While many studies have reported changes in localized activity of the brain in response to memory load [7]–[9], its effect on connectional links among brain sub-networks remains largely unexplored. Nevertheless, a number of recent studies have explored modulation of brain connectivity with load. The between-region interdependence of blood-oxygen-level dependent (BOLD) signals showed a negative correlation with memory load [10] in all parts of the network. However, using DCM and correlation analysis, connectivity was found to enhance and inhibit different connections at different loads [11], [12]. These findings suggested that the functional connectivity strength across the brain is modulated by memory load. However, it should be noted that all the discussed findings are confined to connectivity within lower-frequency bands (<0.1 Hz) whereas, the structure of the functional connectivity network does not remain unchanged within different time scales [13].

Correspondingly, we hypothesize that the TE can be used to investigate the level of interactions or, in other words, the information flow rate between different areas of the brain. Furthermore, the neural pathways that convey information during *encoding* and help to maintain it during the *maintenance* stage are expected to show different levels of interaction with changing memory load. To elucidate these relationships, we examined changes in the information flow among different areas of brain with changing memory load during a working memory task. In order to do so, we first identified the areas of brain where activity co-varied with memory load using

This work was partially supported by the Electrical and Computer Engineering Department of the University of Memphis and NSF grant NSF-IIS-0746790.

independent component analysis (ICA) and then examined the information flow between each pair using the transfer entropy.

II. METHODS

A. Participants

Fifteen graduate students (8 female) participated in the study. Participants were between 24 and 33 years of age (μ±σ: 28 ± 3 yrs). All participants had normal or corrected-to-normal vision and were strongly right-handed as measured by the Edinburgh Handedness Inventory (laterality index > 95%) [14]. Subjects reported no history of visual or neuropsychiatric disorders or were currently on medication. The experiment was undertaken with the understanding and written, informed consent of each participant in compliance with a protocol approved by the University of Memphis Institutional Review Board. Participants were compensated for their time.

B. Experiment

We adopted a modified version of the Sternberg memory task [15]. This task is suitable for studying working memory because it can systematically be configured for different memory loads. Comparing to other WM tasks like n-back task, it also temporally separates encoding, maintenance, and recall stages of the WM process. On each trial, subjects briefly (500 ms) observed a matrix consisting of different English characters positioned around a center point ("SET"; Fig. 1). Characters were displayed with white color over a black background. The size of each character was 1.15°; they were distributed around a center fixation cross and within a visual angle of 2.9°. Array size varied randomly on each trial (2, 4, 6, or 8 items). In all variations of the task, characters were displayed in an array such that their distribution on the left and right side of the center point were the same. After a 3s delay (i.e., *maintenance* stage), a "TEST" character was shown in the center of the screen. Subjects responded via a button press to indicate if this character had occurred in the previous memory "SET".

On half the trials, the test item occurred in the set; the other half it did not. Subjects were encouraged to respond as accurately as possible, and feedback was given via a colored light on the screen, 300 ms after the participant's response. The next trial was initiated after a 3.4 sec inter-stimulus interval. Following 20 practice trials for task familiarization, subjects completed 60 experimental trials per set-size condition.

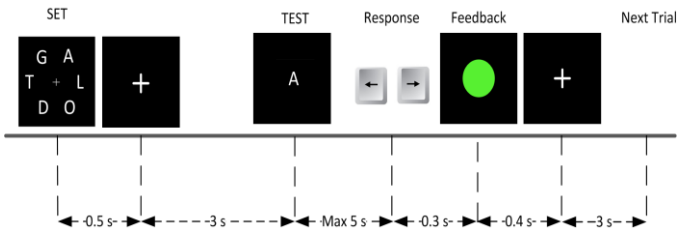

Fig. 1. Time-course of the WM experiment. An array of characters (SET) was displayed for 500ms which followed by a blank screen for 3000ms during which participants were asked to retain the information. Next, a test character (TEST) was shown on the screen and participants responded by pressing a button. Feedback was also shown on the screen to indicate the (in)correct response.

Subjects were seated inside an electro-acoustically shielded booth. They were instructed to avoid body movement and restrict their visual gaze during the task by fixating on the center of the screen. The visual WM task was presented on an LCD monitor at a distance of 1 m. Periodic breaks (~5 min) were given between experimental blocks, which lasted ~15-16 min depending on response speed. The visual stimuli were implemented in MALTAB using the Psychophysics Toolbox [16]. In addition to accuracy, response times (RTs) were also recorded during the experiment, computed as the time-lapse between the appearance of the "TEST" character and the participant's response.

C. EEG Recording

Neuroelectric responses were recorded using standard procedures reported by our laboratory [17], [18]. Briefly, the continuous EEG was recorded from 64 sintered Ag/AgCl electrodes placed around the scalp at standard 10-10 locations [19] (Neuroscan, Quik-cap). Electrodes placed on the outer canthi of the eyes and the superior and inferior orbit were used to monitor ocular activity. Data were digitized with a sampling rate of 500 Hz using an online filter pass-band from DC-250 Hz. Electrode impedance was maintained ≤ 5kΩ over the course of the experiment. During online acquisition, neural responses were referenced to an electrode placed ~1cm posterior to Cz. However, data were re-referenced off-line to a common average reference (CAR) for subsequent analyses.

For the analysis, EEG data were down-sampled to 250 Hz, and base-line corrected by removing the average of each channel. Ocular artifacts (saccades and blink artifacts) were corrected in the EEG using principal component analysis (PCA) [20]. Responses were then filtered using a band-pass filter using a zero phase FIR filter with cut-off frequencies on 1 and 45 Hz. For each set size condition, EEG data was segmented in periods of 9000 ms starting from 2000 ms before presentation of "SET" to 3500 ms post presentation of the test character "TEST". Independent component analysis (ICA) (see part D) was then applied to the set of multi-channel epoched trials, decomposing the recorded signals into statistically independent source signals [21]. The number of independent components was set equal to the number of channels.

Event-related spectral perturbation (ERSP) of each ICA signal was computed to study the time-frequency changes in the EEG across memory loads [22]. Only correct response trials were considered in the analysis. For the current study, the ERSP is desirable as it also captures non-phase-locked neural activity, induced by the stimulus presentation that is not observable with traditional evoked potential averaging [23]. ERSPs were computed by calculating the mean change in spectral power (in dB) from baseline for different frequency and latencies using a Morlet wavelet transform [24]. This analysis achieved a time-frequency resolution of 40 ms/0.5 Hz. The baseline power spectrum was calculated for a 2 sec reference period before the stimulus presentation. The EEGLAB toolbox was used to compute ERSP response [24]. We defined spectral perturbations as the mean power change within each frequency band of interest. Each individual's alpha (α) (8-13) Hz frequency band power was measured from their ERSP. Permutation test ($N = 2000$, $p < 0.001$) was used to

identify time-frequency points of the ERSP which showed significant changes in spectral power (null hypothesis of 0 dB perturbation) relative to the pre-stimulus baseline period.

D. Independent Component Analysis

Neuroelectric signals recorded at each scalp electrode are formed by the summation of different overlapping potentials originating from various brain sources. Estimating coherence among electrodes may lead to spurious inferences on connectivity due to volume conductance. ICA performs linear spatial filtering on the EEG data to isolate independent neuronal sources contributing to the neurophysiological signal recorded at the scalp. In addition, ICA was used to factorize the data into temporally independent components and to create dipolar scalp maps without including any geometrical information about the head or electrode placements. ICA provides a powerful means to isolate brain signals that index physiologically distinct processes [24]–[28]. In the current study, ICA allowed us to identify distinct brain sources that contribute to WM processing which have otherwise been blurred in traditional ERP studies [29].

Group-wise ICA decomposition was used in this study to identify the working memory related effects common across all subjects [6]. ICA was applied on the data set consisting of the correct trials from all four conditions and fifteen participants. Independent components were found using the extended infomax algorithm [30]. Projection vectors corresponding to each independent component (IC) were then extracted from the mixing matrix (W^{-1}) and were used to localize an equivalent dipole [31]. Electrode positions were registered to an MNI (Montreal Neurological Institute) head model template and two symmetric dipoles were fit to each component (one in each hemisphere) using the boundary element method [32] (v12.0.2.5b). Source dipole localization was computed using the DIPFIT plugin in EEGLAB performed on the ICA weighting matrix. Furthermore, only ICs whose sources could be accurately localized and fell inside the head boundary were considered further. Specifically, precision of fitted dipole was measured by assessing the residual variance (RV) of the scalp map of the best-fitting dipole; ICs with RV more than 10% were discarded. Seventeen ICs with RV < 10% were kept from the original set of 64 components.

E. Transfer Entropy (TE)

Transfer entropy is a measure of effective connectivity based on information theory which was proposed by Schreiber [33]. It is a non-parametric method that does not assume any model and detects nonlinear interactions. Using the concept of Markov processes, TE was defined based on the deviation from generalized Markov property in equation (1).

$$p(x_{t+\tau} \mid x_t^{(n)}) = p(x_{t+\tau} \mid x_t^{(n)}, y_t^{(m)}) \quad (1)$$

where $x_t^{(n)} = \{x_t, ..., x_{t-n+1}\}$ and $y_t^{(m)} = \{y_t, ..., y_{t-m+1}\}$ are the realizations of two stochastic processes representing two interacting systems. $x_{t+\tau}$ is the value of process x evaluated at time $t+\tau$. Equation (2) holds only if the the future values of x are independent of y. For instance in the current study, for reasons discussed in the results section, the phase of the ICs were used as the time series x and y. Kullback entropy was used to quantify the level of deviation from equation (1) which leads to definition of "transfer entropy":

$$T_{X\to Y} = \sum p(y_{t+\tau}, y_t^{(m)}, x_t^{(n)}) \log \frac{p(y_{t+\tau} \mid y_t^{(m)}, x_t^{(n)})}{p(y_{t+\tau}, y_t^{(n)})} \quad (2)$$

It was shown by Palus et al. [34] that transfer entropy can be reformulated as a conditional mutual information equation:

$$T_{X\to Y} = I(x_t^{(n)}; y_{t+\tau}^{(m)} \mid y_t^{(n)}) = H(x_t^{(n)}, y_t^{(m)}) + H(y_{t+\tau}^{(m)}, y_t^{(m)}) - H(x_t^{(n)}, y_t^{(m)}, y_{t+\tau}^{(m)},) - H(y_t^{(m)}) \quad (3)$$

Intuitively, this value equals the amount of information in the future of y about the signal x. Nonetheless, unlike mutual information, TE is asymmetric and thus can be used to distinguish causality.

In order to calculate TE, the probability density needs to be computed for the signal. Prior to density estimation, the signal must first be quantized. A key factor in this procedure is the bin size of the quantization that plays a critical role on the performance of the algorithm. The optimal bin size should be selected based on distribution of data which result in a relatively uniform distribution of sample points in all the bins. For instance selecting too small size causes the distribution of sample points to become too sparse and a too large value will lead to overly populated bins. We selected the bin size equal to the optimal bandwidth of a Gaussian kernel density estimator [35] which is given by

$$h = 1.06\sigma N^{-1/5} \quad (4)$$

where σ is the sample standard deviation and N is the number of samples. Conditional mutual information was computed using the MIToolbox [36].

III. Results

A. Selecting the Independent Components

As reported in our previous study [37], ERSP spectrograms were computed for each of the remaining components. It was observed that the dominant activity in the frequency domain was centered on the frequency of 10 Hz. This activity consisted of a transient decrease in α-power during the *encoding* and an increase during *maintenance* (Fig. 2). A clear difference in duration of α-power negativity following the presentation of "SET" was observed across the set sizes which could not be quantified using a fixed-length time window. Therefore, the *encoding* stage was selected as the period starting from disappearance of the stimulus until the time the power decayed back to ¼ of its maximum negativity. Accordingly, the *maintenance* stage started immediately after the *encoding* period and extending to the presentation of "TEST" character. Since we were interested in exploring the changes in

information flow between areas that exhibited a load effect, we only kept those ICs whose response was significantly modulated across set sizes.

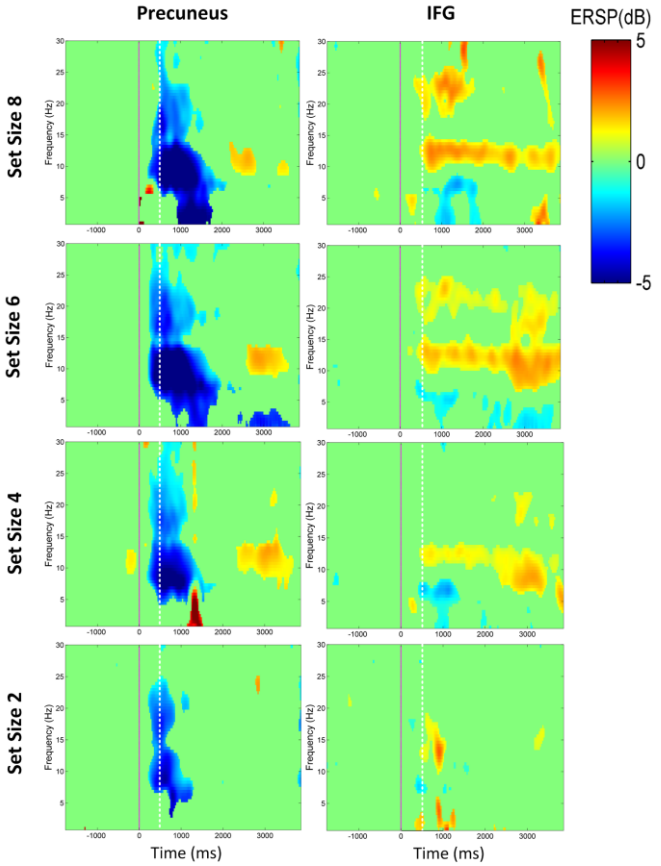

Fig. 2. ERSP spectrogram for activity in Precuneus and IFG as a function of memory load; Hot and cool colors represent increase and decrease in spectral power from baseline, respectively. Time zero and the white dashed vertical lines mark the beginning and end of the presentation of "SET" respectively.

Based on our preliminary observations, the α-power was selected as the response measure for each IC. This choice also corroborated with previous EEG studies which reported load-dependent activity in α-power [8], [38]. Consequently, only four of the components showed significant changes as revealed by a repeated measures ANOVA (Tukey-Kramer adjustments for multiple comparisons) at an a priori α-level of $p = 0.05$. *Encoding* α-power in bilateral precuneus ($F_{3,36} = 5.86$, $p = 0.002$), cuneus ($F_{3,36} = 5.54$, $p = 0.003$), and fusiform gyrus (FG) ($F_{3,36} = 3.17$, $p = 0.035$) and *maintenance* α-power in bilateral precuneus ($F_{3,36} = 3.49$, $p < 0.03$) and inferior frontal gyrus (IFG) ($F_{3,36} = 2.51$, $p = 0.07$) showed significant changes and were selected for the subsequent TE connectivity analysis.

Fig. 3 shows the α-power observed in each region across the four set sizes. The average *encoding* α-power decreased with increasing set size from size 2 to 6 and plateaued thereafter. In contrast, average *maintenance* α-power increased from set size 2 to 4 and leveled out after. Table I summarizes the location and corresponding residual variance of each of these IC components.

TABLE I. SOURCE LOCATIONS OF SELECT INDEPENDENT COMPONENTS

IC#	Talairach Coord. (x,y,z)	Location	Closest Brodmann Area (BA)	RV%
1	-8,-55,37 8,-55,37	Bilateral Precuneus	7	4.74
2	-47,7,24 47,7,24	Bilateral Inferior Frontal Gyrus	9	5.43
3	-45,-59,-11 45,-59,-11	Bilateral Fusiform Gyrus	37	3.42
4	-19,-77,34 19,-77,34	Bilateral Cuneus	7	2.40

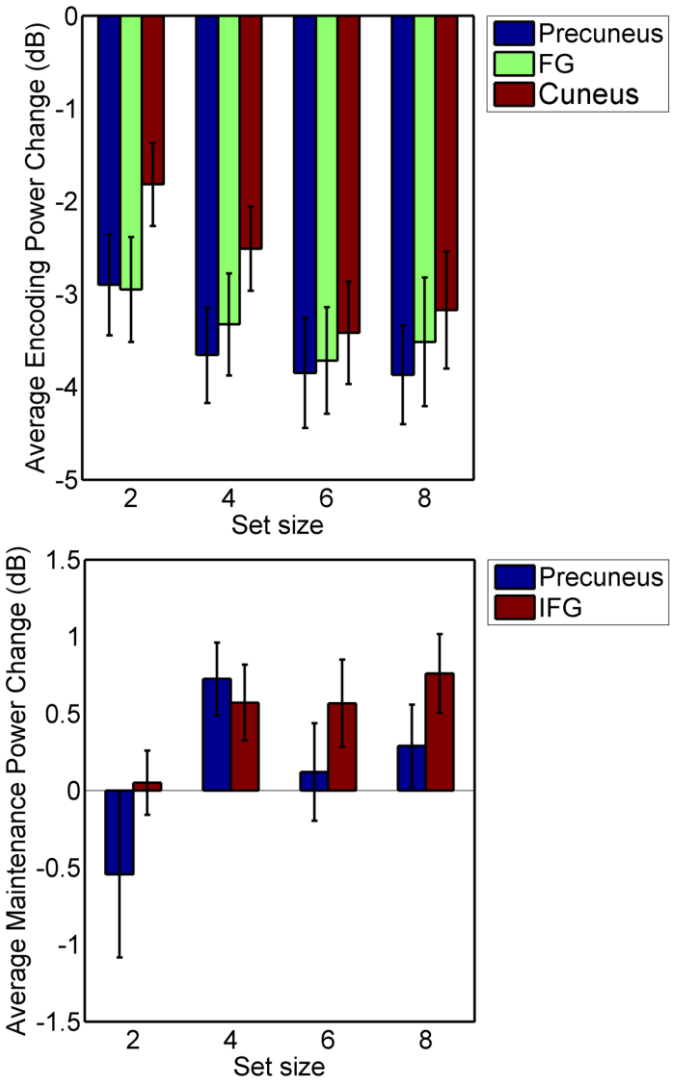

Fig. 3. Difference in α-band power in select regions of brain during encoding (top) and maintenance (bottom) stages.

B. Computing the Transfer Entropy

Previous studies have shown that even if the amplitudes of two coupled chaotic oscillators remain uncorrelated, their phases may synchronize [39]. Hence in this study, instead of raw activity signals, IC phases were computed using the Hilbert transform and were used to compute TE values. Computation of phase for a signal with no well-expressed peak

in its power spectrum can lead to spurious findings [40] and hence IC activity signals were filtered within the α-frequency band (8-13Hz) with a band-pass zero phase FIR filter (n=500) before applying the Hilbert transform. The reason for choosing the α-band activity for connectivity analysis was the well-established role of the neural activity within this frequency band in working memory [8], [38].

Transfer entropy was used to detect nonlinear interdependencies between pairs of IC phase signals. We computed the value of TE during *baseline*, *encoding* and *maintenance* periods (see *Results* part *A*). In order to directly compare TE values across conditions and periods (*baseline*, *encoding* and *maintenance*), an equal number of points needed to be sampled from each. Therefore, we chose a one second window from each of the periods as below:

1) *Baseline window:* starting 1000 ms before presentation of "SET"
2) *Encoding window:* starting from 300 ms and extending up to 1300 ms after presentation of "SET". The 300 ms delay was considered to exclude any possible delays in brain response after the presentation of stimulus (see Fig. 2).
3) *Maintenance window:* starting 1000 ms before presentation of "TEST"

In addition, TE values were computed for a range of lags ($\tau = 1...15$) and then averaged across lags. This procedure was employed to decrease the variance of estimation error and to avoid detecting spurious high values of interdependency [6]. Moreover, in order to study the changes in information transfer rate initiated by the cognitive task it was necessary to compute TE values relative to the baseline. Therefore, the relative transfer entropy (rTE) for *encoding* and *maintenance* stages were computed by subtracting the baseline TE value from each.

$$rTE_{\underset{enc}{X\to Y}} = TE_{\underset{enc}{X\to Y}} - TE_{\underset{base}{X\to Y}}$$
$$rTE_{\underset{ma\,int}{X\to Y}} = TE_{\underset{ma\,int}{X\to Y}} - TE_{\underset{base}{X\to Y}} \quad (5)$$

TABLE II. SIGNIFICANT TEST RESULTS OF CHANGES IN RELATIVE TE ACROSS SET SIZES (X = *NONSIGNFICANT*)

Encoding Period				
Source IC	***Destination IC***			
	1	2	3	4
1		0.002	X	X
2	X		X	X
3	0.006	0.006		X
4	0.06	0.02	X	
Maintenance Period				
Source IC	***Destination IC***			
	1	2	3	4
1		X	X	0.006
2	X		X	0.004
3	0.04	X		0.003
4	X	X	X	

rTE values were computed for each IC pair and all trials and were then averaged over trials for each individual and set size. Furthermore, in order to identify the specific links between IC components whose information transfer rates were modulated by memory load, we applied a repeated measures ANOVA test on TE values for each IC pair. The p-values from these tests are summarized in Table II.

A directed graph can be utilized to visualize the connections across the engaged brain regions [1]. The nodes on this graph are the brain regions represented by the location of the fitted dipoles and likewise, edges correspond to the links between the ICs which reached significance according to the values from Table II. The direction of the link is determined by the specific pair and its corresponding ANOVA. Fig. 4 illustrates the directed graph for connectivity during the *encoding* stage. *rTE* values for three of the links including the links from nodes (i.e., ICs) 1, 3 and 4 to node 2 decreased with increasing memory load while the links from nodes 3 and 4 to node 1 increased mainly from set size 4 to 6. Likewise, Fig. 5 shows a similar graph for the connectivity during the *maintenance* stage. Comparatively, all the links in this graph showed a decreasing trend in *rTE* which declined with increasing memory load level.

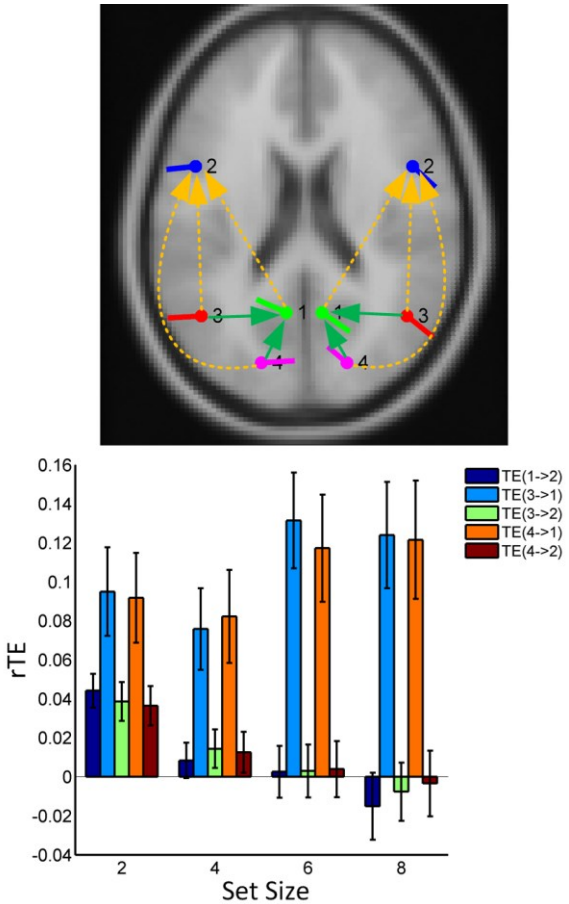

Fig. 4. WM Encoding Period Connectivity; (top) graph of load dependent connectivity; the positive and negative changing edges are shown by solid green and dashed orange arrows, respectively. (bottom) rTE values for the graph edges across the four set sizes; node (i.e., IC) 1: precuneus, node 2: IFG, node 3: FG, node 4: cuneus;

IV. Discussion

We examined the degree of directed interaction between different brain areas whose responsiveness co-varied with memory load. We applied group-wise ICA to decompose the multi-channel scalp EEG into a set of temporally independent components. The set of components whose activations co-varied with memory load was selected for subsequent connectivity analysis. Interdependencies between these select regions of the brain were investigated using transfer entropy. Possible neural pathways conveying stimulus/memory information between these areas during *encoding* and *maintenance* stages of the visual memory task were identified.

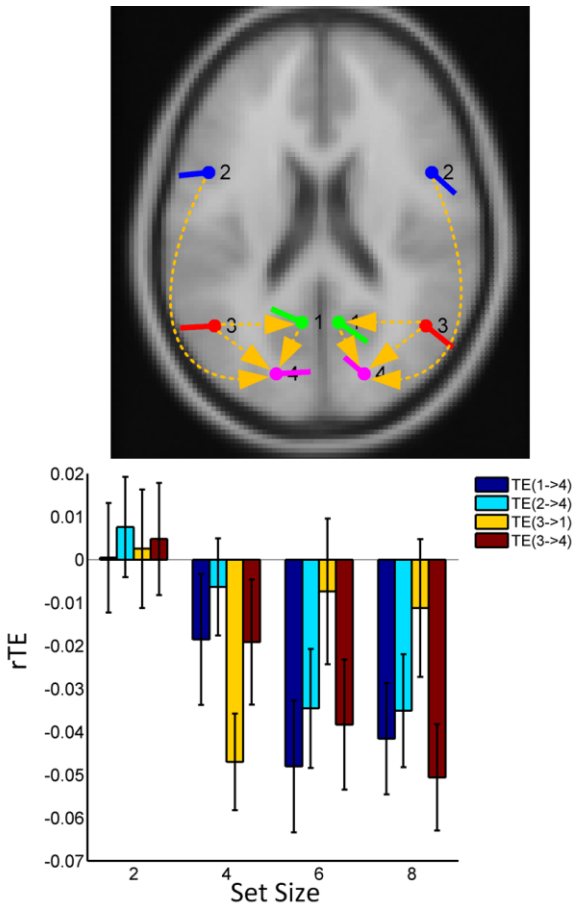

Fig. 5. WM Maintenance Period Connectivity; (top) graph of load dependent connectivity; negative changing edges are shown by dashed orange arrows (bottom) rTE values for the graph edges across the four set sizes; node (IC) 1: precuneus, node 2: IFG, node 3: FG, node 4: cuneus;

Results obtained from connectivity analysis showed that during the *encoding* stage, the information transfer rate within the pathways connecting posterior to anterior brain regions were modulated by the memory load (Fig. 4). However, the graph corresponding to the *maintenance* stage connectivity exhibited a reversal in the directionality of this neural pathway, resulting in modulation of prominent information flow from anterior to posterior regions (Fig. 5). Judging by the dominant direction of the graph edges within the encoding and *maintenance* periods, these results are generally in line with the current knowledge about the role of occipital areas (node 4) in processing visual stimulus information, parietal areas (node 1) in visual working memory and storage, and frontal areas (node 2) in executive function and control. The reversed direction of the graph during the *maintenance* stage also aligns with the notion of the reverse hierarchy theory (RHT) [41]. RHT posits that with increasing task demand, the brain performs a progressive backward search from higher- order information to lower-level inputs in search of representations with optimal signal-to-noise ratio. It is possible that the reversal of TE directionality we observe (cf. Figs. 4 and 5) reflects a similar cascade of "top-down" guided modulation where task-relevant visual information must be enhanced at the input (occipital cortex) and irrelevant information pruned before being broadcast to higher brain areas for storage (temporal-parietal cortices; IC1) and memory retrieval (frontal cortices; IC2).

Moreover, during the *encoding* stage, *rTE* values increased with increasing memory load in pathways connecting cuneus and FG to precuneus region (3 -> 1) and (4 -> 1) whereas, it decreased for the other links. While this implies an increase in the information transfer in the former it also suggests the opposite pattern between IFG and other regions. On the other hand, during the *maintenance* stage, *rTE* values for all the links were negatively affected by the increasing memory load. Interestingly, the decreasing trend of TE with increasing memory load suggests a weakening in between-region connections which could also be interpreted as inhibition of cross-talk between areas [42]. This observation is corroborated by a previous connectivity study on fMRI data in which the mean between-region correlations were reported to significantly decrease with increasing memory load [10].

V. Conclusion

This paper introduces an approach to quantify the level of interaction among different brain regions. Empirical analyses revealed that information flow during the *encoding* and *maintenance* stages follows different pathways in opposite direction. This approach could potentially be used in identifying sub-networks of brain and their interactions under different conditions. This in turn could lead to building a more comprehensive model of the working brain to better explain the patterns of neural activity related to specific cognitive tasks or those altered by neurological disorders.

References

[1] E. Bullmore and O. Sporns, "Complex brain networks: graph theoretical analysis of structural and functional systems.," Nat. Rev. Neurosci., vol. 10, no. 3, pp. 186–98, Mar. 2009.

[2] P. Fries, "A mechanism for cognitive dynamics: neuronal communication through neuronal coherence.," Trends Cogn. Sci., vol. 9, no. 10, pp. 474–80, Oct. 2005.

[3] V. Sakkalis, "Review of advanced techniques for the estimation of brain connectivity measured with EEG/MEG.," Comput. Biol. Med., vol. 41, no. 12, pp. 1110–7, Dec. 2011.

[4] R. Vicente, M. Wibral, M. Lindner, and G. Pipa, "Transfer entropy--a model-free measure of effective connectivity for the neurosciences.," J. Comput. Neurosci., vol. 30, no. 1, pp. 45–67, Feb. 2011.

[5] M. Wibral, B. Rahm, M. Rieder, M. Lindner, R. Vicente, and J. Kaiser, "Transfer entropy in magnetoencephalographic data: quantifying

information flow in cortical and cerebellar networks.," Prog. Biophys. Mol. Biol., vol. 105, no. 1–2, pp. 80–97, Mar. 2011.

[6] V. A. Vakorin, N. Kovacevic, and A. R. McIntosh, "Exploring transient transfer entropy based on a group-wise ICA decomposition of EEG data.," Neuroimage, vol. 49, no. 2, pp. 1593–600, Jan. 2010.

[7] P. Bashivan, G. M. Bidelman, and M. Yeasin, "Neural correlates of visual working memory load through unsupervised spatial filtering of EEG," in Proceedings of 3rd workshop on Machine Learning and interpretation in neuroimaging, 2013.

[8] O. Jensen, J. Gelfand, J. Kounios, and J. E. Lisman, "Oscillations in the alpha band (9-12 Hz) increase with memory load during retention in a short-term memory task.," Cereb. Cortex, vol. 12, no. 8, pp. 877–82, Aug. 2002.

[9] J. D. Cohen, W. M. Peristein, T. S. Braver, L. E. Nystrom, D. C. Nolls, J. Jonides, and E. E. Smith, "Temporal dynamics of brain activation during a working," Nature, vol. 386, no. April, 1997.

[10] C. E. Ginestet and A. Simmons, "Statistical parametric network analysis of functional connectivity dynamics during a working memory task.," Neuroimage, vol. 55, no. 2, pp. 688–704, Mar. 2011.

[11] L. Ma, J. L. Steinberg, K. M. Hasan, P. A. Narayana, L. A. Kramer, F. G. Moeller, and A. Larry, "Working memory load modulation of parieto-frontal connections: Evidence from dynamic causal modeling," Hum. Brain Mapp., vol. 33, no. 8, pp. 1850–1867, 2012.

[12] J. Rissman, A. Gazzaley, and M. D'Esposito, "Dynamic adjustments in prefrontal, hippocampal, and inferior temporal interactions with increasing visual working memory load.," Cereb. Cortex, vol. 18, no. 7, pp. 1618–29, Jul. 2008.

[13] C. J. Honey, R. Kötter, M. Breakspear, O. Sporns, and R. Ko, "Network structure of cerebral cortex shapes functional connectivity on multiple time scales.," Proc. Natl. Acad. Sci. U. S. A., vol. 104, no. 24, pp. 10240–5, Jun. 2007.

[14] R. C. Oldfield, "The assessment and analysis of handedness: the Edinburgh inventory," Neuropsychologia, vol. 9, no. 1, pp. 97–113, 1971.

[15] S. Sternberg, "High-speed scanning in human memory," Science (80-.)., vol. 153, no. 3736, pp. 652–654, 1966.

[16] D. H. Brainard, "The psychophysics toolbox," Spat. Vis., vol. 10, no. 4, pp. 433–436, 1997.

[17] G. M. Bidelman, J. W. Villafuerte, S. Moreno, and C. Alain, "Age-related changes in the subcortical-cortical encoding and categorical perception of speech.," Neurobiol. Aging, pp. 1–15, May 2014.

[18] G. M. Bidelman, S. Moreno, and C. Alain, "Tracing the emergence of categorical speech perception in the human auditory system.," Neuroimage, vol. 79, pp. 201–12, Oct. 2013.

[19] R. Oostenveld and P. Praamstra, "The five percent electrode system for high-resolution EEG and ERP measurements," Clin. Neurophysiol., vol. 112, no. 4, pp. 713–719, 2001.

[20] G. L. Wallstrom, R. E. Kass, A. Miller, J. F. Cohn, and N. A. Fox, "Automatic correction of ocular artifacts in the EEG: a comparison of regression-based and component-based methods," Int. J. Psychophysiol., vol. 53, no. 2, pp. 105–119, 2004.

[21] S. Makeig, A. J. Bell, T.-P. Jung, and T. J. Sejnowski, "Independent component analysis of electroencephalographic data," Adv. Neural Inf. Process. Syst., pp. 145–151, 1996.

[22] S. Makeig, "Auditory event-related dynamics of the EEG spectrum and effects of exposure to tones," Electroencephalogr. Clin. Neurophysiol., vol. 86, no. 4, pp. 283–293, 1993.

[23] D. L. Schomer and F. L. Da Silva, Niedermeyer's electroencephalography: Basic principles, clinical applications, and related fields. Wolters Kluwer Health, 2012.

[24] S. Makeig, S. Debener, J. Onton, and A. Delorme, "Mining event-related brain dynamics," Trends Cogn. Sci., vol. 8, no. 5, pp. 204–210, 2004.

[25] T. P. Jung, S. Makeig, M. Westerfield, J. Townsend, E. Courchesne, and T. J. Sejnowski, "Analysis and visualization of single-trial event-related potentials.," Hum. Brain Mapp., vol. 14, no. 3, pp. 166–85, Nov. 2001.

[26] A. C. Tang, B. a Pearlmutter, N. a Malaszenko, D. B. Phung, and B. C. Reeb, "Independent components of magnetoencephalography: localization.," Neural Comput., vol. 14, no. 8, pp. 1827–58, Aug. 2002.

[27] A. Lenartowicz, A. Delorme, P. D. Walshaw, A. L. Cho, R. M. Bilder, J. J. McGough, J. T. McCracken, S. Makeig, and S. K. Loo, "Electroencephalography correlates of spatial working memory deficits in attention-deficit/hyperactivity disorder: vigilance, encoding, and maintenance.," J. Neurosci., vol. 34, no. 4, pp. 1171–82, Jan. 2014.

[28] R. Vigário, J. Särelä, V. Jousmäki, M. Hämäläinen, and E. Oja, "Independent component approach to the analysis of EEG and MEG recordings.," IEEE Trans. Biomed. Eng., vol. 47, no. 5, pp. 589–93, May 2000.

[29] E. K. Vogel and M. G. Machizawa, "Neural activity predicts individual differences in visual working memory capacity," Nature, vol. 428, no. April, pp. 748–751, 2004.

[30] T. W. Lee, M. Girolami, and T. J. Sejnowski, "Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources.," Neural Comput., vol. 11, no. 2, pp. 417–41, Feb. 1999.

[31] T. F. Oostendorp and A. Van Oosterom, "Source parameter estimation in inhomogeneous volume conductors of arbitrary shape," Biomed. Eng. IEEE Trans., vol. 36, no. 3, pp. 382–391, 1989.

[32] A. Delorme and S. Makeig, "EEGLAB : an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," vol. 134, pp. 9–21, 2004.

[33] T. Schreiber, "Measuring information transfer," Phys. Rev. Lett., vol. 85, no. 2, p. 461, 2000.

[34] M. Palus and M. Vejmelka, "Directionality of coupling from bivariate time series: how to avoid false causalities and missed connections.," Phys. Rev. E. Stat. Nonlin. Soft Matter Phys., vol. 75, no. 5 Pt 2, p. 056211, May 2007.

[35] B. W. Silverman, Density estimation for statistics and data analysis, vol. 26. CRC press, 1986.

[36] G. Brown and M. Luj, "Conditional Likelihood Maximisation : A Unifying Framework for Information Theoretic Feature Selection," J. Mach. Learn. Res., vol. 13, no. 1, pp. 27–66, 2012.

[37] P. Bashivan, G. M. Bidelman, and M. Yeasin, "Spectrotemporal dynamics of the EEG during working memory encoding and maintenance predicts individual behavioral capacity," Eur. J. Neurosci., vol. In press, 2014.

[38] A. M. Tuladhar, N. ter Huurne, J.-M. Schoffelen, E. Maris, R. Oostenveld, and O. Jensen, "Parieto-occipital sources account for the increase in alpha activity with working memory load.," Hum. Brain Mapp., vol. 28, no. 8, pp. 785–792, Aug. 2007.

[39] A. Pikovsky, M. Rosenblum, J. Kurths, and R. C. Hilborn, "Synchronization: a universal concept in nonlinear science," Am. J. Phys., vol. 70, no. 6, p. 655, 2002.

[40] M. G. Rosenblum, A. S. Pikovsky, and J. Kurths, "Synchronization approach to analysis of biological systems," Fluct. noise Lett., vol. 4, no. 01, pp. L53–L62, 2004.

[41] M. Ahissar and S. Hochstein, "The reverse hierarchy theory of visual perceptual learning.," Trends Cogn. Sci., vol. 8, no. 10, pp. 457–64, Oct. 2004.

[42] O. Jensen and A. Mazaheri, "Shaping functional architecture by oscillatory alpha activity: gating by inhibition.," Front. Hum. Neurosci., vol. 4, no. November, p. 186, Jan. 2010.