Neurophysiological and Behavioral Correlates of Human-Multiagent Task Performance

Gregory Bales *Student Member, IEEE,* and Zhaodan Kong *Member, IEEE*

***Abstract*—An effective human-robot team will incorporate the cognitive skills of the human with the autonomous capabilities of the robot group to maximize task performance. However, producing a seamless fusion will require a greater understanding of the complex cognitive state of the human as it reacts to inherent uncertainties in both the task environment and robot dynamics. In this paper, we investigate how variations in neurophysiological and behavioral measures correlate with performance in a human-multiagent team task. This study utilizes external behaviors in concert with spectral power and functional connectivity, acquired via electroencephalography (EEG), to probe the intricate interactions between cognitive processes, behaviors, and task performance. We show that characteristic changes in the α (8-12Hz) and θ (4-8Hz) bands of EEG, indicate a higher burden on the cognitive resources associated with visual-spatial reasoning that occur during decreases in task performance. These results were reinforced by complementary behavioral shifts in both the gaze and piloting inputs of the human subjects. Finally, higher performing subjects tended to engage more actively in the task, utilizing a greater amount of visual-spatial reasoning, to perform more effectively.**

***Index Terms*—Electroencephalogram, functional connectivity, human-multiagent performance.**

I. Introduction

HUMAN-robot teams are expected to provide solutions in a wide range of applications, such as environmental monitoring, human directed search and rescue, or hazard containment and mobilization. These groups may consist of tens of robotic agents that coordinate their actions to achieve a common goal. Compared with one, or several agents working independently, a robotic group holds the advantage of being more efficient at covering larger areas. These groups also have the benefit of scalability: individual agents can be added or removed without significantly affecting the performance of the group as a whole [1], [2]. In many applications [3] humans are required to play an active rather than supervisory role. Such types of human-multiagent tasks give rise to a fundamental problem: is it possible to optimize a natural and effective collaboration between the group and its human partner? Considering the various cognitive constraints placed upon the human, this problem can be particularly difficult to solve.

Manuscript received... This work was supported by a Space Technology Research Institutes grant (grant number 80NSSC19K1052) from NASA's Space Technology Research Grants Program. *(Corresponding author: Zhaodan Kong)*

G. Bales and Z. Kong are with the Department of Mechanical and Aerospace Engineering, University of California, Davis, CA, 95616 (e-mail: zdkong@ucdavis.edu)

Among the variety of challenges that need to be addressed are those related to: 1) the cognitive load of the human operator; 2) the communication method of the robotic group; and 3) the interface between the human-robot team, e.g. methods by which the human can control the agents. Researchers have made key advancements by testing human subjects within simulated environments [4], [5]. Such methods are justified by the fact that controlled environments greatly limit the effects of confounding factors. However, it has been observed [6] that given identical tasks, testing in real, rather than simulated environments can increase the workload and reaction times [7] of the human subjects.

A major goal in human-robot interaction research is a fusion between the autonomous capabilities of the robot(s) and the cognitive skills of the human, while maximizing task performance, efficiency, and the simplicity of the interaction. Achieving this goal requires examining human factors, such as trust [8], choice [9], and limited attention [10]. Additionally, we must consider suitable levels of autonomy, control sharing [11], and aspects of human behavior and cognition in the interaction design. Significant work has attempted to organize networked systems to influence the behavior of a multirobot team [12], the characteristics of a multiagent group [13], or to construct an effective interaction framework [14]. Other work has examined what information can be effectively passed to the human operator [15] or how its presentation can affect the interaction [16]. However, these efforts do not examine detailed aspects of human cognitive behavior. The ACT-R cognitive architecture has been utilized to serve as a proxy for the cognitive state of the operator that was used to model choice behavior [17]. However, the results were highly contingent upon model fidelity. There is still a need to examine neurophysiological measures in real time that can provide insight into the complex cognitive state of the human.

One important application of human-robot teaming is robot-assisted urban search and rescue (RUSaR) [19], which can be highly visual-spatial in nature. Successful task performance may require acute levels of situational awareness that can impose a heavy burden on the operator's attention system [7]. Attention is an essential property of cognitive operation that can be roughly divided into two categories [20]: 1) bottom-up, a stimulus-driven process in which a salient sensory event captures our attention; and 2) top-down, a voluntary goal-driven process based on aspects, such as tasks, knowledge, expectations, and memory. The attention system is controlled by a distributed network of brain regions [21]. Signals acquired via electroencephalography (EEG) have been used to examine memory performance [22] and cognitive workload [23]. EEG

has been employed in social human-robot interaction research [24] and Brain Computer Interface (BCI) implementations. [25]. Recently, EEG has also been used to examine fatigue and drowsiness [26] in aircraft pilots and to identify periods of distraction during vehicle driving [27]. These studies primarily include the magnitude and spatial distribution of spectral power in the θ (4-8Hz) and α (8-12Hz) bands. Theta oscillations located in the mid frontal region of the brain are relevant to working memory upkeep [28]. By comparison, alpha band oscillations over the entire head occur in all the historic cognitive domains of perception, attention, and access to working memory. Functional long range α band interactions exert top down influences to inhibit sensory susceptibility. Brain regions that are activated during a task exhibit a decrease in α power, whereas task irrelevant regions show an increase in α power. Therefore, α oscillations can serve to gate attention [29].

The combination of neurophysiological with behavioral characteristics can depict a more comprehensive mental state: they may indicate how human's cognitive processes might affect their behaviors that ultimately determine performance, and, furthermore, how these relationships, shown in Fig. 1, transpire. Test scenarios in which human subjects interact with real robots provide the opportunity to gain insight into how humans handle, not only the complexities of robot dynamics and variations in environment, but also the inherent uncertainties in the human-robot interaction.

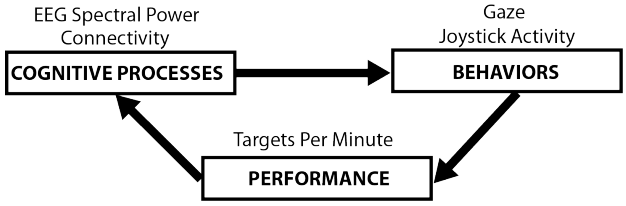

Fig. 1. Performance to Behavior Loop connecting the cognitive processes that drive behaviors, which in turn affect the overall task performance. This figure includes the five variables utilized in this study.

Given the attention-demanding nature of many human-robot team tasks, a natural question arises: *Does the* ***geometric complexity (GC), defined as the number of agents required to estimate the kinematic state of the robotic group****, impact task performance, and if so, is this reflected in behavioral and neurophysiological measures?* It has been shown [31] that the performance of a human-multiagent team task will decrease with an increasingly complex spatial distribution of the agents. A decline in performance would be a natural consequence of the increase in mental computation [7] necessary to continuously estimate the kinematic state of the group. Therefore, these cognitive demands should present themselves in the α and θ band features of EEG. In addition, measurable changes in external behaviors, such as gaze and pilot input activity, should also be present. These results would be indicative of the visual-spatial nature of the task.

Our previous works were based on data collected from a series of human subject experiments. In our first conference paper [30], we provided a general description of these experiments and presented a preliminary analysis based on gaze patterns and cognitive load using established methodologies. In our second conference paper [31], we examined the EEG data as a collective measure using established spectral power analysis. In addition, we showed that task performance decreases with increasing GC. The major contributions of this paper are as follows: 1) We offer the most comprehensive analysis of the data we have collected by investigating human-multi-agent interaction as cognition affects behaviors which determine performance, as shown in Fig. 1; 2) We propose a network-cognitive-scientific approach of analyzing EEG data. As far as we know, this is one of the first instances such a methodology has been applied to analyze human multi-agent performance using data obtained from testing in a real, rather than simulated environments; 3) We show that the decrease in task performance associated with GC is directly reflected in neurophysiological and behavioral measures. Furthermore, we show that higher performing subjects engage more actively in the task by utilizing a greater amount of visual-spatial reasoning. The neurophysiological and behavioral measurements in this study are obtained in real time. Consequently, our results may contribute to future developments in the areas of augmented cognition [32] and human-swarm interaction [1].

II. Experiment

Section I defined geometric complexity (GC) as the number of agents required to estimate the kinematic state of the robotic group. In order to determine the impact of GC on human-robot performance, we conducted a series of human subject experiments inspired by simulation studies [18] but with real ground robots and humans locally embedded in the task space. These tests were based on a target identification and acquisition task. We utilized an EEG device to measure changes in both spectral power and functional connectivity in the α and θ bands that are associated with cognitive activity. In addition, a pair of eye-tracking glasses was used to measure gaze behavior. The results obtained from this study are derived from the full complement of neurophysiological and behavioral measures listed in Table I.

Section II-A describes our experimental setup. Details of the individual tests are defined in Section II-B. Finally, our experimental hypotheses are outlined in Section II-C.

A. Experimental Setup

The study consisted of a 4.2 m by 5.5 m test arena surrounded by motion tracking cameras. The test arena will be referred to as the "task space". This task space contained ground robots under the control of locally embedded human subjects. Two wall mounted projectors displayed interactive targets onto the floor.

We have integrated several modules into a measurement suite to record behavioral, neurophysiological, and mechanical data. Each module is named for its primary measure: 1) Gaze: a wearable eye tracking system by SensoMotoric Instruments; 2) EEG: an Emotiv Epoc headset [33] with an array of 14 electrodes positioned via the International 10-20 system; and 3) 3D Pose: an Optitrack motion capture system consisting of 12 wall mounted cameras. A detailed description of the Gaze

and 3D Pose modules, along with their processing methods were developed [30], [31]. These modules are shown in Fig. 2, along with one of the target-to-target trials in this paper.

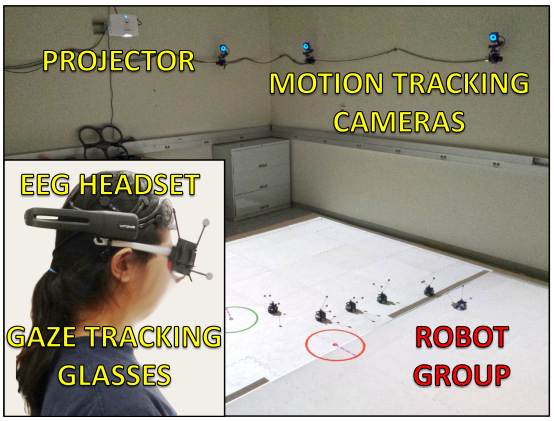

Fig. 2. The human-robot interaction arena with motion tracking cameras and projectors, as well as a robot group, and a subject outfitted with gaze tracking glasses.

The coordinate system established by the Optitrack module will be referred to as the "world" frame, denoted by W. In addition, each trial was recorded using an overhead camera. An example is shown in Fig. 4. Piloting commands of linear and angular velocity (v and ω respectively) from the human subjects were applied with a Logitech Force 3D Joystick. The commands were sampled at 30Hz and smoothed with a first-order Butterworth filter at a corner frequency of 3Hz.

B. Test Trials

Our test subjects were required to pilot a group of six Pololu m3pi differential drive vehicles in each of two configurations. Only the input from a single joystick was used to pilot the group. In order to simplify control, reverse commands were not allowed. A centralized controller calculated the six independent non-interacting path-following control laws [34], with simple minimum inter-robot spacing rules to prevent collisions. Maximum speed limits ensured stability. The control signals were broadcast wirelessly to each individual robot.

The two multirobot configurations are shown in Fig. 3. For the serpentine configuration, the human subject would steer a lead cart while the five trailing robots would simply follow the path established by the leader. The rectangular configuration was piloted by steering the central motion of a virtual rigid body. The six unique command paths were established by the motion of virtual carts positioned about the center of a rectangular virtual shape, as seen in Fig. 3. The controlled motion of the rectangular configuration deviated from its perfect shape depending upon the aggressiveness of the human subject's steering commands. Examples of the group motion can be found at https://youtu.be/QoLUWKFrHWA and https://youtu.be/TpCnO3kb2jo. The deviation of the carts from a perfect rectangular shape is also visible in Fig. 4 and Fig. 5. Detailed illustrations of the controller behavior can be found at http://cphslab.com/publications.

The kinematics of the robotic group can be described mathematically by the motion of vectors in a given coordinate system. We denote the position of cart i in the world frame as r_i^W. The origin of serpentine configuration o_{srp}^W was defined by the position of the third robot in the chain: $o_{srp}^W = r_3^W$. We chose this position to distribute each subject's focus over the larger group, rather than the lead robot. In contrast, the origin of the rectangular configuration was defined as the instantaneous centroid of the six carts: $o_{rct}^W = \frac{1}{6}\sum_{i=1}^{6} r_i^W$. Piloting this group required human subjects to estimate the true position of the centroid.

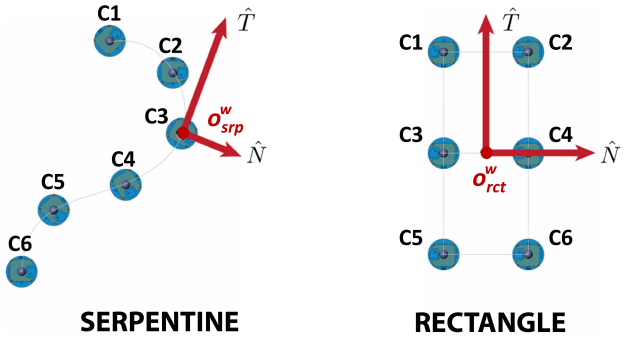

Fig. 3. The two multirobot configurations. The tangential $\hat{T}$ and normal $\hat{N}$ directions of the body frame for each configuration are labeled. The origin for the serpentine configuration o_{srp}^W is defined by position of the third robot in the chain, whereas the origin of the rectangular configuration o_{rct}^W is defined by its centroid.

During each trial, subjects were presented with targets projected onto the floor. Each target contained heading information indicated by a line. Subjects were tasked to pilot the origin of each robotic group from one target to the next as quickly as possible. New targets were only revealed when both a minimum origin-to-target distance of 15.2 cm (6 in) and a minimum heading difference of $\pm 15°$were met. Therefore, only two targets, a start and a finish, were visible to the subjects at any time. A detail of a target pair is shown in Fig. 4. All target-to-target paths were designed to be feasible given the spatial constraints of the task space. Subjects were provided eight minutes to pilot each configuration with the explicit task of acquiring as many targets as possible. Depending on the skill of the human subject, the number of targets acquired ranged from 4 to 12. There was a five minute rest phase between each trial.

Using the lead robot, a subject could effectively acquire targets with the serpentine configuration by tracking, at the very least, two robots: the lead cart; and the third cart. However, the rectangular configuration may require all six robots for adequate estimation of the origin depending on the distortion of the controlled shape. Consequently, we assigned the serpentine a GC of 2, and the rectangle a GC of 6. Regardless of how aggressively the subject attempts to pilot the rectangular configuration, we expect it to be more difficult to perform the task effectively than the serpentine configuration.

Ten subjects completed the tests. Each subject was an undergraduate student between the ages of 18 to 22. All subjects were instructed on both the testing procedure and proper use of the test equipment. In addition, there were two training sessions in which subjects would control first a single robot, then all six robots in each configuration. We

TABLE I
DETAIL OF THE DEPENDENT VARIABLES UTILIZED IN THIS STUDY

VARIABLE	DESCRIPTION
Targets Acquired per Minute: T_{pm}	Task purpose, and the direct metric of performance. Total number of successful targets obtained in an 8 minute trial: Section II-C
Spatial Gaze Distribution: G_F, G_M	Gaze distribution in the *Forward* and *Middle* regions of the body frame of the robotic configuration: Section III-A
Joystick Activity: J_A	Cumulative sum of incremental joystick movement: Section III-B
EEG Power: $\hat{P}_\alpha, \hat{P}_\theta$	Normalized spectral power in the α and θ bands: Section III-E
Regional Connectivity: C_α, C_θ	Pairwise phase consistency derived from the complex coherency in the α and θ bands: Section III-G

demonstrated for them how to acquire the target using each configuration. For all subjects, the test series proceeded as follows: training to the serpentine trials; and finally to the rectangular trials.

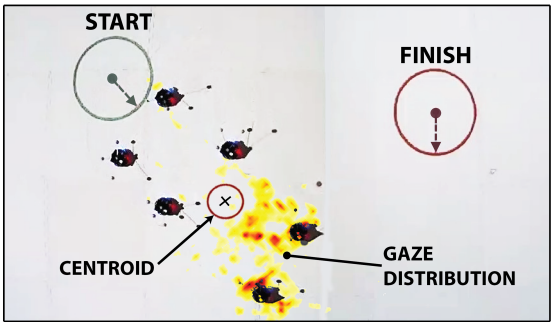

Fig. 4. An overhead screen capture of a single target attempt. Both the position and heading information of the target are displayed. A subject must pilot the configuration of robots from an initial target (START) to a final target (FINISH). Note that the gaze distribution, indicated by a heat map, leads the true centroid.

We examine the arrival phase: the period during which a subject attempts the acquisition of a target. This phase encompasses the final eight seconds before the configuration reaches the target, as detailed in Fig. 5. We chose not to discriminate between a successful and an unsuccessful target acquisition. The subjects are generally unaware of their outcome until after the configuration has passed over the target point. Therefore, regardless of success or failure, each subject intends to succeed and behaves in a manner to do so.

C. Experimental Hypotheses

The explicit objective of our human-robot tasks was to acquire the maximum number of targets in a fixed period of time. Consequently, targets acquired per minute (T_{pm}) is the inherent metric of task performance that is employed in this study. In [31] we showed that task performance decreased by 48% as a consequence of increasing the GC. In this study, we developed two hypothesis that further explore the correlations between the reduction in task performance and neurophysiological and behavioral characteristics. The dependent variables d are briefly outlined in Table I and their specific details are given in Section III. The experimental hypotheses are:

1) Hypothesis 1: The greater cognitive resources that are required to pilot the rectangular configuration will present measurable differences in the average neurological and behavioral measures. More specifically, θ power in the mid and frontal regions and α power in the occipital regions of the brain should increase. In addition, long range functional connectivity in the α band would indicate top down control of the visual attention process. These outcomes would imply an increase in visual-spatial attention and cognitive processing, all of which are consistent with well-established cognitive scientific literature. Furthermore, measurable differences in gaze and pilot input activity should be present. The study was within-subject, and the single factor Configuration had two levels: "serpentine" and "rectangle".

2) Hypothesis 2: The subjects with higher performance have a natural proclivity for this particular human-robot task. Therefore, *they will utilize fewer cognitive resources than those with lower performance*. The lower cognitive demands should be reflected in the average neurological and behavioral measures, as described in *Hypothesis 1*. The study was between-subject, and the single factor Performance had 2 levels: "high" and "low".

III. DATA ANALYSIS

In this section we discuss the methods used to prepare our data. Two primary pipelines were established to generate data for the analysis that follows: one to extract the gaze data; and the other for the EEG spectral powers and functional connectivity. Data streams were filtered and temporally aligned for comparison. The description of these pipelines are provided in Sections III-A and III-C.

A. Gaze Data

1) Rotation to the Body Frame: A world frame W for the task space was established by the Optitrack motion capture system. A pose estimate of the subject's head was combined with the gaze, extracted from the eye tracking glasses, to synthesize the gaze vector in W as r_g^W. The intersection of this gaze vector with the arena floor created a heatmap of the overall distribution of gaze points, as shown in Fig. 5.

All data were transformed into a body centered coordinate system B, determined for each configuration. In doing so, we could perform an analysis that was independent from the pose of the robotic group. The system B was defined as follows: first, extract the unit vector $\hat{T}$, originating from the rigid body origin o_c^W and tangential to the direction of movement. Next,

project a unit vector $\hat{N}$ perpendicular to $\hat{T}$. Finally, translate by o_c^W and multiply by the rotation matrix R_W^B from the world frame W to the body frame B to obtain

$$r_g^B = R_W^B(r_g^W - o_c^W), \tag{1}$$

where r_g^B is the gaze in the body frame. We transformed the cart positions, target positions, and the gaze distributions into the new body frame B. An example of these transformations for a single target attempt is shown in Fig. 5.

2) Estimating Gaze Distributions: The sample probability of gaze within a given region was determined as the sum of all gaze points r_g^B falling in a particular region η divided by the N total gaze samples contained in each 8 s target period. We denote the estimate of this probability as G_η:

$$G_\eta = \frac{1}{N}\sum_{\eta} r_g^B. \tag{2}$$

One such distribution is also illustrated in Fig. 5.

The wider dispersion of carts in the rectangular configuration naturally induces increased gaze activity in the direction of $\hat{N}$. We examined differences in gaze distributions along the $\hat{T}$ direction. Two areas were selected for comparison: the middle region, denoted by $\eta = M$ and defined between $\pm$20.3 cm from the centroid; and the forward region, denoted by $\eta = F$ and defined as >20.3 cm from the centroid.

B. Joystick Activity

In order to quantify pilot inputs from each human subject, we defined the joystick activity J_A as a unitless measure. The J_A for the N samples in each 8 s target period was calculated as the cumulative sum of incremental differences in normalized joystick commands for the linear, $\Delta\hat{v}_n = [v(t_n) - v(t_{n-1})]/v_{max}$, and angular, $\Delta\hat{\omega}_n = [\omega(t_n) - \omega(t_{n-1})]/\omega_{max}$, velocities by

$$J_A = \sum_{n=1}^{N}\sqrt{\Delta\hat{v}_n^2 + \Delta\hat{\omega}_n^2}. \tag{3}$$

C. EEG Spectral and Connectivity Data

The EEG data reduction pipeline is outlined in Fig. 6. Many of these pipeline methods are commonly used for feature identification and extraction in BCI implementations [25]. Data reduction was performed as follows:

1) Record EEG data for a single trial attempt.
2) Filter the data with a zero lag, 4th-order bandpass filter (2-20Hz) and prune to 8 s segments preceding the acquisition of each target.
3) Perform an independent component analysis to identify and remove spurious artifacts. Reconstruct the 14 channel EEG from the remaining independent components.
4) Separate the data into 1 s epochs, offset at 50 ms intervals.
5) At each epoch, calculate the power spectral densities in the individual bands, α and θ. Determine the per channel spectral power at the central frequencies of each band using (4). Extract the normalized power using (5). Average each channel over the 8 s period.
6) At each epoch, calculate the Pairwise Phase Consistency (PPC) in the individual bands, α and θ, between each channel pair. Average each pair over the 8 s period.

Finally, in order to facilitate statistical testing, we reduced the full 14 node graph into a compact 4 node graph using (7).

D. Artifact Removal Using Independent Component Analysis

It has been shown that independent component analysis is an effective means of identifying statistically independent neurological sources [35]. We chose the FastICA algorithm [36] to generate the independent components. Before proceeding with the analysis, components associated with electromyographic artifacts arising from blinks and lateral eye movements were removed. No more than two components were identified for removal in any data set.

E. Spectral Power

The power spectral density $S_{f_j}^i$ (units of $\mu V^2/Hz$) for each EEG channel i was calculated for the j discrete frequencies f_j using a fast Fourier transform at a spectral resolution of Δf. The spectra were determined over 1 s epochs offset at increments of ΔT = 50 ms. This yielded temporal sequences $S_{f_j}^i(t_l)$ at each time epoch $l \in \{1, 2, \ldots, L\}$ spanning the target period, such that $L\Delta T$ = 8 s. Next, we numerically integrated the power spectral sequences over the frequencies in each band $\nu \in \{\alpha, \theta\}$ to obtain power in μV^2 as

$$P_\nu^i(t_l) = \Delta f \sum_{f_j \in \nu} S_{f_j}^i(t_l). \tag{4}$$

The powers were normalized using baseline data by

$$\hat{P}_\nu^i(t_l) = \frac{P_\nu^i(t_l) - B_\nu^i}{B_\nu^i}, \tag{5}$$

where B_ν^i was obtained from the human subject at the start of each trial. Finally, we calculated the target spectral powers P_α^i and P_θ^i as the mean of each temporal sequence over the 8 s target period.

F. Principal Component Transformation of Spectral Powers

The full set of EEG data were arranged into a matrix $X \in \mathbb{R}^{14\times p}$ for the p target observations accumulated by all 10 subjects. These data exhibited mild linear correlations (R^2= 0 to 0.5) between individual channels. A valid statistical test required the estimation and removal of these correlations. However, the potential interactions between a full complement of 14 variables can make this intractable in practice. We removed the linear interrelationships by rotating the data matrix into a coordinate frame, defined by principal components that were determined via diagonaliztion of the covariance matrix: $\Sigma = (X - \bar{X})(X - \bar{X})^T$. The 14 unit eigenvectors w_i of Σ formed a basis set for the spatial distribution of spectral power. These principal components were arranged to form the columns of an orthonormal matrix $W \in \mathbb{R}^{14\times 14}$. Multiplying the mean centered EEG data by W produced a new, nearly uncorrelated ($R^2 < 10^{-6}$) data set K, expressed as $K = W(X - \bar{X})$.

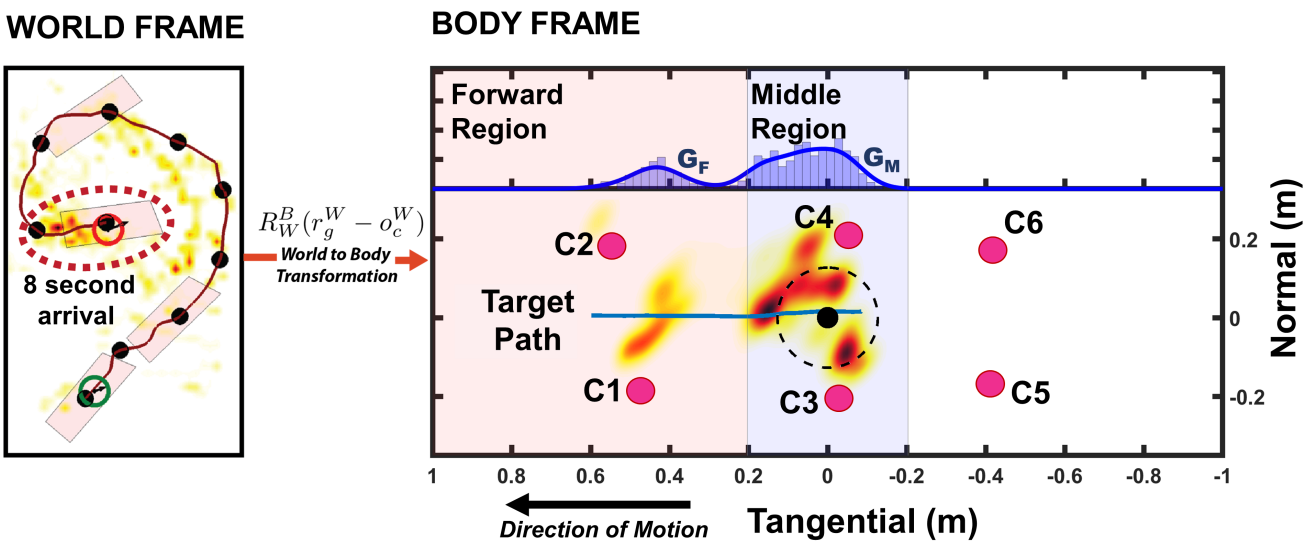

Fig. 5. Detail of gaze distributions for a rectangular configuration. Gaze data, shown as a heat map, are projected onto the arena floor. These data are transformed from the world frame W to the moving body frame B of the robotic group using (1). The gaze regions and the associated histogram are delimited by vertical lines. Each of the gaze probabilities, G_F and G_M, are defined in (2). Average cart positions are marked as $C1$ through $C6$. The group origin (centroid) along with the 30.5 cm diameter threshold region is shown by a black dot and circle, respectively. Note the deviation of the cart positions from a true rectangular distribution.

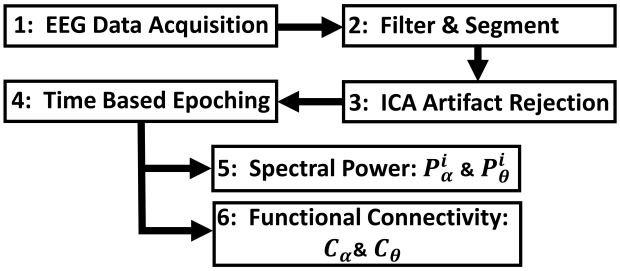

Fig. 6. Detail of six steps used in the EEG data reduction.

Our statistical testing proceeded on the data K using an analysis of variance (ANOVA) between the factors described in Section II-C. By enumerating the total set of principal components as $\mathcal{W} = \{1, 2, \cdots, 14\}$ and the subset of K whose differences were statistically significant be given as $\mathcal{A} \subseteq \mathcal{W}$, we reconstructed differences in the spatial distribution of power X^{B-A} between the factors A and B as

$$X^{B-A} = \sum_{i \in \mathcal{A}} w_i (\bar{k}_i^B - \bar{k}_i^A), \tag{6}$$

where w_i is the principal component, and $\bar{k}_i$ is the average value of k_i extracted from the ANOVA. Only components with $p < 0.05$ were selected for comparison.

G. Functional Connectivity

There are a wide array of functional connectivity measures that are commonly use in EEG analysis [37]. We selected the PPC, which lies between 0 and 1, and has the advantage of greater statistical power [38] by taking an average of the $\binom{N}{2}$ binary combinations of phase angles in N ensembles, rather than the N angles themselves. We calculated the PPC values by extracting the relative phase between channels i and j, determined from an estimate of the complex coherency for each of the bands, α and θ. In the remainder of this paper, references to connectivity between channels specifically refers to the PPC values between them.

1) Regional Connectivity: Functional connectivity of the 14 EEG channels (AF3, F3, FC5, F7, T7, P7, O1, O2, P8, T8, F8, FC6, F4, AF4) can be represented by a fully connected undirected weighted graph, as shown in the top left headmap of Fig. 7. This graph possesses an adjacency matrix $A^F \in \mathbb{R}^{14\times14}$, where A_{ij}^F is the PPC value between nodes i and j. Communication between brain regions is indicated by PPC values close to unity [39]. However, thresholding our data near values of unity yielded extremely sparse graphs, an example is shown in the bottom left headmap of Fig. 7. Moreover, the likelihood of any specific connection existing among all the subjects was very low. Consequently, the ANOVA failed to yield statistical differences among subject groups. However, the data did indicate that connections between four larger regions followed general trends. These regions were divided by the anterior, posterior, right, and left, as shown in the top left headmap of Fig. 7. In order to extract conclusions from the observed trends, we reduced our graph from 14 channels, to the four regions. We specified a set $\mathcal{R} = \{AR, AL, PR, PL\}$ in which the 14 EEG channels were collected into these four regions:

AR (anterior-right) = {AF4, F4, F8, FC6};
AL (anterior-left) = {AF3, F3, F7, FC5};
PR (posterior-right) = {T8, P8, O2};
PL (posterior-left) = {T7, P7, O1}.

Next, we defined the elements of the adjacency matrix $A^R \in \mathbb{R}^{4\times4}$ to represent the reduced graph between regions. These values were determined by the sum total of PPC measures between EEG channels in $a, b \in \mathcal{R}$, normalized by the total possible connections between them, thereby ensuring that the

reduced connectivity measure would still fall between 0 and 1. If $|a|$ and $|b|$ are set cardinalities, then A^R_{ab} is given by

$$A^R_{ab} = \frac{1}{|a||b|} \sum_{\substack{i \in a, j \in b \\ i \neq j}} A_{ij}. \quad (7)$$

The four node graph reduces the number of unique connections from 91 to ten, as illustrated in the top two headmaps of Fig. 7; however, it expresses the same information as the 14 node graph, but looses the granularity of the individual channels. A comparison between a sparse 14 node graph and its reduced four node equivalent is detailed in bottom two headmaps of Fig. 7. Our analysis proceeded on the reduced connectivity between the four regions.

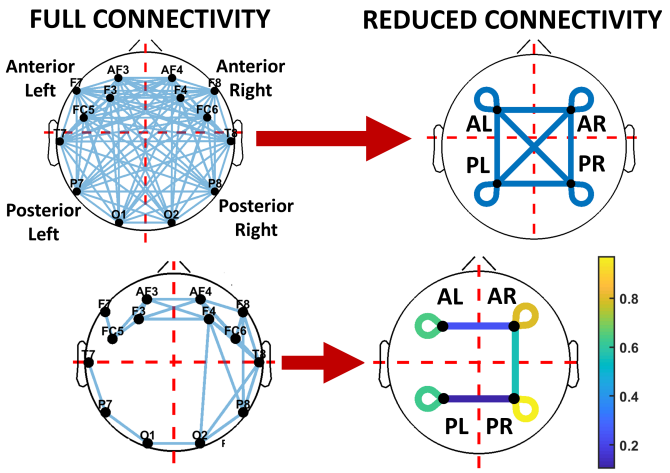

Fig. 7. Detail of connectivity reduction. Combining the connections within the four regions reduces the sparse 14×14 adjacency matrix to a compact 4×4 adjacency matrix. The bottom headplots illustrate one example of a sparse 14×14 graph and its reduced equivalent.

IV. Results and Discussion

The tests for all subjects proceeded as follows: from training; to the serpentine trials; and finally the rectangular trials. If any transferable learning occurred, it would have been from the serpentine to the rectangular configuration. The training phase was designed to provide the subjects with ample experience piloting both configurations. We note that the absence of counterbalance may be a potential confound that can limit an interpretation of the results.

Section IV-B presents the results of the within-subject factor of Configuration with levels of "serpentine" and "rectangle". Section IV-C presents the results for the between-subject factor of Performance with levels of "high" and "low". High and low performance values were selected based on the average number of targets per minute between both configurations as derived from our previous work [31]. A threshold value of 1.1 targets per minute effectively divided the population into two groups of five higher, and five lower performing subjects. These threshold values are given in Table II.

A. Statistical Analysis

We performed all statistical tests as single factor ANOVAs with Bonferroni post hoc correction. The dependent variables

TABLE II
High and Low Performance Thresholds

Performance	$\mathbf{T_{pm}}$ Threshold
HIGH performing subjects	$T_{pm} > 1.08$
LOW performing subjects	$0 < T_{pm} \leq 1.08$

d are outlined in Table I. It was assumed under the null hypothesis that the distribution of d between factors A and B was identical, i.e., H_0: $\bar{d}_A = \bar{d}_B$. There existed a mild correlation between gaze distributions of the middle and forward regions ($R^2 = 0.5$). We fit a linear least-squares model between these two variables, and the ANOVA proceeded using their residuals.

B. Hypothesis 1: Impact of Geometric Complexity on Neurophysiological and Behavioral Measures

This section examines the the ANOVA for the within-subject factor of Configuration. These data represent the measures of the serpentine configuration subtracted from those of the rectangular configuration ($\Delta = rectangle - serpentine$).

The statistically significant differences in EEG characteristics are summarized in Fig. 8. The results for the α band are displayed in the left column, while those for the θ band are displayed in the right column. Differences in normalized band power, defined in (5), for each of the 14 EEG channels are displayed in the top plots. The red shaded area represents a 95% confidence interval on the means. For reference, head regions are noted according to the groups defined in Section III-G1. Each headmap illustrates the spatial distribution of the same data presented in the plot directly above it. Differences in the regional connectivity, defined in (7), are indicated in the lower graph.

The statistically significant differences in behavioral data are summarized in Fig. 9. Error bars represent the 95% confidence interval on the means.

1) Spectral Power and Connectivity: Fig. 8 shows that $\Delta\hat{P}_\alpha$ was a minimum of -4% in the right and left frontal regions. The $\Delta\hat{P}_\alpha$ then increased to a maximum of 3% at O2. These results illustrate that piloting the serpentine configuration resulted in greater power in the frontal region, and less power in the occipital, temporal, and parietal regions.

Fig. 8 also shows that there was a 21% increase in α connectivity between the anterior-right to posterior-right regions.

The right and left frontal regions displayed positive values in $\Delta\hat{P}_\theta$, from 15 to 60%, indicating more power in those regions associated with the rectangular configuration. Furthermore, the left temporal region showed negative values of $\Delta\hat{P}_\theta$, as low as -100%, indicating larger serpentine related power.

There are no statistically significant differences in connectivity between regions in the θ band.

2) Behavioral Characteristics: In Fig. 9 we see that subjects spent 16% more time looking in the middle region of the rectangular configuration. In addition, joystick activity decreased by 24%.

3) Discussion: Examining the differences, we see that piloting the rectangular configuration produced several key results. The higher θ power in the frontal region indicates more access to working memory resources. Furthermore, the greater

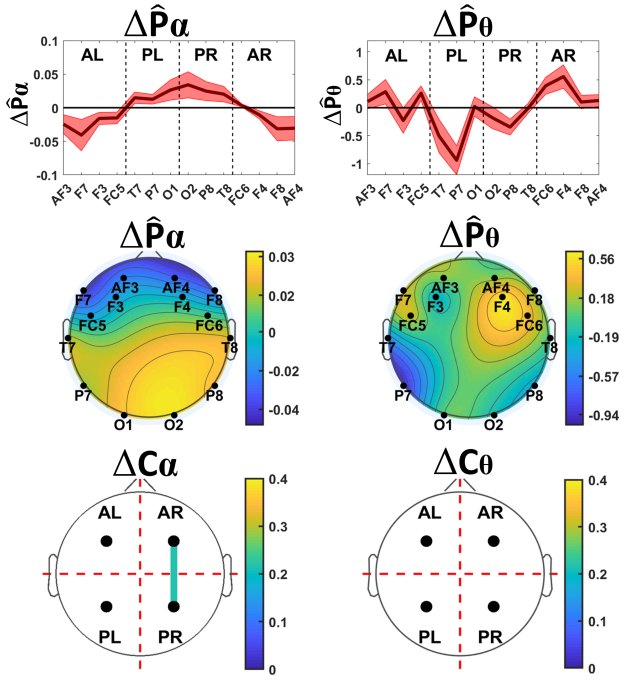

Fig. 8. Comparison of the neurophysiological differences associated with the factor Configuration with $\Delta = rectangle - serpentine$. These are the results of the within-subject, single factor ANOVA with Bonferroni post hoc correction ($p < 0.05$). Shaded areas represent the 95% confidence interval on the means.

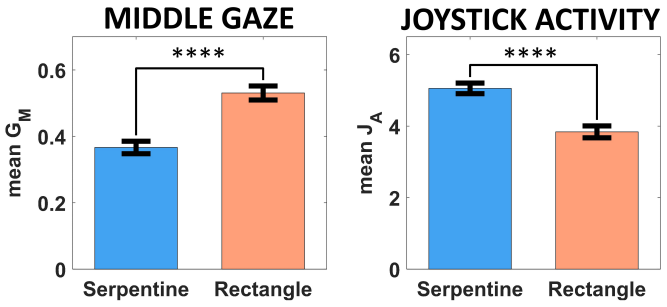

Fig. 9. Comparison of the behavioral differences associated with the factor Configuration. Error bars represent upper and lower estimates of the means.

α power in the occipital region is associated with the suppression of task irrelevant visual stimuli. Additionally, the increase in α connectivity between the anterior-right to posterior-right regions is also indicative of top-down suppression of visual stimuli. This α connectivity may further imply a greater focus on internal attention [40]. Finally, subjects spend more time looking in the middle region of the rectangular configuration, yet they utilized less pilot input.

We can obtain insight into these results by comparing the methods for piloting the configurations. Each subject controls the serpentine configuration by steering only the lead cart; the remaining robots follow the path established by the leader. Both the reduction of gaze in the middle region and the greater amount of joystick activity are behaviors that result from the confident and predictably deterministic control of the single lead robot. In contrast, subjects pilot the rectangular configuration by steering the collective motion of a virtual rigid body. Rather than a single cart, each subject must comprehend the relatively fluid motion of all six robots at once. A determination of the centroid requires constant visual estimation. The focused attention that is necessary to accomplish this results in larger amount of gaze in the middle region and the utilization of more working memory resources. Additionally, the indirect interface may also explain the reduction in control activity, since the subject's internal map, from pilot input to configuration motion, requires a greater amount of mental computation to establish a similar degree of confidence. However, there is also the possibility that gross positioning of the robotic group may have already occurred and that little positional fine tuning occurred. Regardless, the suppression of visual stimuli indicates a larger focus on internal cognitive processes. From a cognitive point of view, there is simply more internal processing required to pilot the rectangular configuration. The outcomes are generally slower speeds and more missed targets. These results are directly consistent with *Hypothesis 1*.

C. Hypothesis 2: Characteristics of High/Low Performers

In this section we examine the results of the ANOVA on the between-subject factor of Performance. Data are given as the measures of lower performing subjects subtracted from those of the higher performing subjects ($\Delta = high - low$). The statistically significant differences in EEG characteristics are summarized in Fig. 10. The statistically significant differences in behavioral data are summarized in Fig. 11. Error bars represent the 95% confidence interval on the means.

1) Spectral Power and Connectivity: Fig. 10 shows that $\Delta\hat{P}_\alpha$ at channels F7 and AF4 was positive, while the remainder in the frontal region was negative. Similarly, $\Delta\hat{P}_\alpha$ at channels T8 and O1 was negative, while the remainder in the anterior region was positive.

For higher performing subjects, α connectivity increased 19% from the anterior-right to posterior-right regions. There were also positive differences in the anterior-left to anterior-right (18%), and the posterior-left to posterior-right (6%) regions. Additionally, there was a 28% increase in the within-posterior-left region.

With the exception of F3 and F8, $\Delta\hat{P}_\theta$ was positive over the frontal region, from 0 to 100%, indicating that high performers exhibit more power in that region.

The increases in θ connectivity include: 25% from the anterior-right to anterior-left regions; 25% in the within-anterior-right and within-anterior-left regions; 31% from the anterior-right to posterior-right regions; and 39% in the within-posterior-left region.

2) Behavioral Characteristics: Shown in Fig. 11, high performing subjects spent 1.7% more time looking in the forward region. These subjects also had a 35% increase in joystick activity.

3) Discussion: Examining these differences, we see that the distribution of α power is complex, showing no real preference for any region of the head. In addition, the literature currently does not strongly support a correlation between connectivity in the θ band and visual-spatial reasoning. Such results are generally difficult to interpret, and we will not do so.

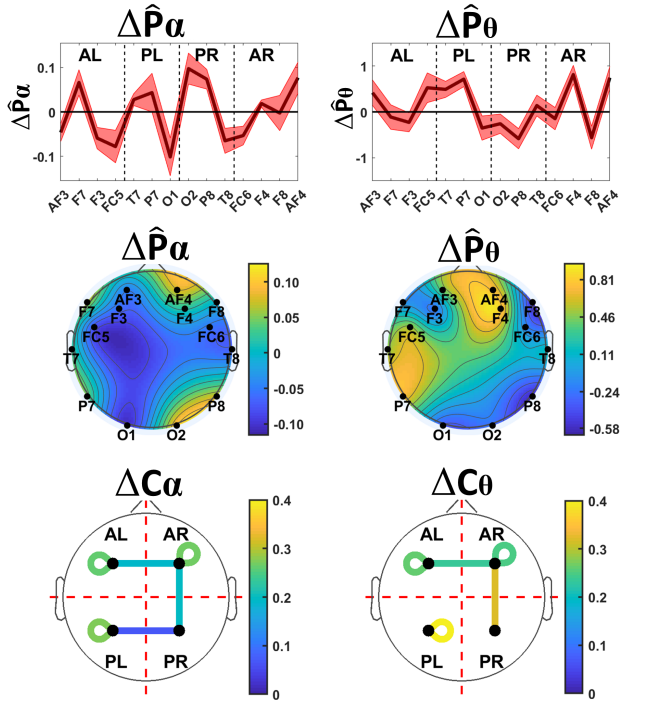

Fig. 10. Comparison of the neurophysiological differences associated with the factor Performance with $\Delta = hi - low$.

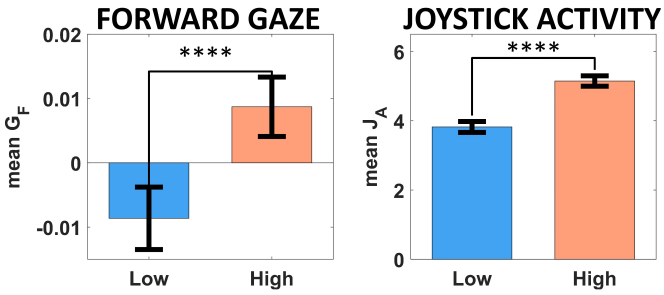

Fig. 11. Comparison of the behavioral differences associated with the factor Performance.

However, there are clear trends in the remaining data. Higher performing subjects had a 19% increase in α connectivity from the anterior-right to posterior-right brain regions, which indicates top-down suppression of visual stimuli. High performers also displayed a 28% increase in within-posterior-left connectivity and an increase in α power at T7 and P7. These changes are indicative of the suppression of sensory information. However, the within-posterior-left result could also be associated with O1 connectivity in the posterior-left region. Unfortunately the reduced connectivity eliminates our ability to discern between the two. Finally, the higher performing subjects exhibited more θ power in the frontal region, which indicates a greater utilization of working memory resources.

Given these observations, the differences between subject groups conforms to a predictable pattern. Higher performing subjects drive the configurations faster and miss fewer targets, but while doing so, they utilize a greater amount of visual-spatial reasoning and internal processing. Additionally, they spend a greater amount of time looking in the forward region, while simultaneously using more control input. However, these results stand in direct contrast to *Hypothesis 2*. Interestingly, higher performing subjects are not naturally more effective with the given control interface. Our results suggest that high performers use more cognitive resources to perform at a higher level, possibly indicating that they are more engaged in the task.

V. Conclusion

This paper examined whether the reduction in human-multiagent team task performance due to an increase in the geometric complexity of a robotic group is reflected in average neurophysiological and behavioral measures. Ten subjects were locally embedded in a task space. Each subject piloted six ground robots in two geometric configurations: a serpentine ($GC = 2$); and a rectangle ($GC = 6$). Our tests revealed that the use of a higher GC configuration yielded a decrease in task performance and that this reduction was accompanied by increases in spectral power and functional connectivity that indicate more internal processing, access to working memory, and suppression of visual stimuli. Complementary changes in gaze and pilot input enforce the conclusion that increasing the GC places a greater burden on the human subject's cognitive resources that detrimentally affect the outcome of a human-multiagent team task. Finally, higher performing subjects, regardless of the GC, tended to engage more actively in the task, utilizing a greater amount of visual-spatial reasoning to perform more effectively. This is one of the first studies to utilize the distribution of EEG spectral power in concert with functional connectivity, gaze, and control input, to examine the complex interaction between cognitive processes, behaviors, and task performance.

References

[1] A. Kolling, P. Walker, N. Chakraborty, K. Sycara, and M. Lewis, "Human interaction with robot swarms: a survey," *IEEE Trans. on Human-Machine Systems*, vol. 46, no. 1, pp. 9–26, 2016.

[2] B. Pendleton and M. Goodrich, "Scalable human interaction with robotic swarms," in *AIAA Infotech@ Aerospace (I@ A) Conf.*, 2013, p. 4731.

[3] S. Bashyal and G. K. Venayagamoorthy, "Human swarm interaction for radiation source search and localization," in *Swarm Intelligence Symposium, 2008.*. IEEE, 2008, pp. 1–8.

[4] A. Kolling, K. Sycara, S. Nunnally, and M. Lewis, "Human swarm interaction: An experimental study of two types of interaction with foraging swarms," *J. of Human-Robot Interaction*, vol. 2, no. 2, pp. 103–128, 2013.

[5] M. L. Cummings and A. S. Brzezinski, "Global vs. local decision support for multiple independent uav schedule management," *Int. J. of Applied Decision Sciences*, vol. 3, no. 3, pp. 188–205, 2010.

[6] G. Podevijn, R. O'Grady, C. Fantini-Hauwel, and M. Dorigo, *Human responses to stimuli produced by robot swarms - the effect of the reality-gap on psychological state*. Springer Int. Publishing, 2018, pp. 531–543.

[7] A. Steinfeld, T. Fong, D. Kaber, M. Lewis, J. Scholtz, A. Schultz, and M. Goodrich, "Common metrics for human-robot interaction," in *Proc. of the 1st ACM SIGCHI/SIGART Conf. on Human-robot interaction - HRI '06*. ACM Press, 2006, p. 33.

[8] L. R. Marusich, J. Z. Bakdash, E. Onal, M. S. Yu, J. Schaffer, J. O'Donovan, T. Höllerer, N. Buchler, and C. Gonzalez, "Effects of information availability on command-and-control decision making: performance, trust, and situation awareness," *Human factors*, vol. 58, no. 2, pp. 301–321, 2016.

[9] L. F. Bertuccelli and M. L. Cummings, "Operator Choice Modeling for Collaborative UAV Visual Search Tasks," *IEEE Trans. on Systems, Man, and Cybernetics - Part A: Systems and Humans*, vol. 42, no. 5, pp. 1088–1099, Sep 2012.

[10] J. W. Crandall, M. L. Cummings, M. Della Penna, and P. M. A. de Jong, "Computing the Effects of Operator Attention Allocation in Human Control of Multiple Robots," *IEEE Trans. on Systems, Man, and Cybernetics - Part A: Systems and Humans*, vol. 41, no. 3, pp. 385–397, May 2011.
[11] S. Musić and S. Hirche, "Control sharing in human-robot team interaction," *Annu. Reviews in Control*, vol. 44, pp. 342–354, 2017.
[12] M. Egerstedt, S. Martini, M. Cao, K. Camlibel, and A. Bicchi, "Interacting with Networks: How Does Structure Relate to Controllability in Single-Leader, Consensus Networks?" *IEEE Control Systems*, vol. 32, no. 4, pp. 66–73, Aug. 2012.
[13] D. S. Brown, S. C. Kerman, and M. A. Goodrich, "Human-swarm interactions based on managing attractors," in *Proc. of the 2014 ACM/IEEE Int. Conf. on Human-robot interaction*. ACM, 2014, pp. 90–97.
[14] J.-P. d. l. Croix and M. Egerstedt, "Analyzing human-swarm interactions using control Lyapunov functions and optimal control," *Networks and Heterogeneous Media*, vol. 10, no. 3, pp. 609–630, Jul. 2015.
[15] A. Li, W. Luo, S. Nagavalli, N. Chakraborty, and K. Sycara, "Handling state uncertainty in distributed information leader selection for robotic swarms," in *Systems, Man, and Cybernetics (SMC)*. IEEE, 2016, pp. 64–69.
[16] P. Walker, M. Lewis, and K. Sycara, "The effect of display type on operator prediction of future swarm states," in *Systems, Man, and Cybernetics (SMC)*. IEEE, 2016, pp. 521–526.
[17] K. Sycara, C. Lebiere, Y. Pei, D. Morrison, and M. Lewis, "Abstraction of analytical models from cognitive models of human control of robotic swarms," in *Int. Conf. on Cogn. Modeling*. University of Pittsburgh, 2015.
[18] P. Walker, S. A. Amraii, N. Chakraborty, M. Lewis, and K. Sycara, "Human control of robot swarms with dynamic leaders," in *Intelligent Robots and Systems (IROS 2014)*. IEEE, 2014, pp. 1108–1113.
[19] Y. Liu and G. Nejat, "Robotic urban search and rescue: a survey from the control perspective," *J. of Intelligent & Robotic Systems*, vol. 72, no. 2, pp. 147–165, Nov. 2013.
[20] M. I. Posner, *Cognitive neuroscience of attention*. Guilford Press, 2011.
[21] S. M. Doesburg, N. Bedo, and L. M. Ward, "Top-down alpha oscillatory network interactions during visuospatial attention orienting," *NeuroImage*, vol. 132, pp. 512–519, May 2016.
[22] W. Klimesch, "Eeg alpha and theta oscillations reflect cognitive and memory performance: a review and analysis," *Brain research reviews*, vol. 29, no. 2, pp. 169–195, 1999.
[23] P. Zarjam, J. Epps, and N. H. Lovell, "Beyond subjective self-rating: Eeg signal classification of cognitive workload," *IEEE Trans. on Autonomous Mental Development*, vol. 7, no. 4, pp. 301–310, 2015.
[24] B. A. Urgen, M. Plank, H. Ishiguro, H. Poizner, and A. P. Saygin, "EEG theta and Mu oscillations during perception of human and robot actions," *Frontiers in Neurorobotics*, vol. 7, 2013.
[25] L. F. Nicolas-Alonso and J. Gomez-Gil, "Brain Computer Interfaces, a Review," *Sensors*, vol. 12, no. 2, pp. 1211–1279, Jan. 2012.
[26] G. Borghini, L. Astolfi, G. Vecchiato, D. Mattia, and F. Babiloni, "Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness," *Neuroscience & Biobehavioral Reviews*, vol. 44, pp. 58–75, Jul. 2014.
[27] Y.-K. Wang, T.-P. Jung, and C.-T. Lin, "Theta and Alpha Oscillations in Attentional Interaction during Distracted Driving," *Frontiers in Behavioral Neuroscience*, vol. 12, Feb. 2018.
[28] F. Roux and P. J. Uhlhaas, "Working memory and neural oscillations: alpha–gamma versus theta–gamma codes for distinct WM information?" *Trends in Cognitive Sciences*, vol. 18, no. 1, pp. 16–25, Jan. 2014.
[29] W. Klimesch, "Alpha-band oscillations, attention, and controlled access to stored information," *Trends in Cognitive Sciences*, vol. 16, no. 12, pp. 606–617, Dec. 2012.
[30] G. Bales and Z. Kong, "Neurophysiological and behavioral studies of human-swarm interaction tasks," in *2017 IEEE Int. Conf. on Systems, Man, and Cybernetics (SMC)*, Oct 2017, pp. 671–676.
[31] G. Bales and Z. Kong, "Cognitive correlates of eeg spectral power indicate human-swarm task performance," in *Int. Workshop on Human-In-The-Loop Internet of Things Systems (HiL-IoT)*, Oct. 2018, pp. 1–6.
[32] D. Schmorrow, K. M. Stanney, G. Wilson, and P. Young, "Augmented cognition in human–system interaction," *Handbook of Human Factors and Ergonomics, 3rd Edition*, pp. 1364–1383, 2006.
[33] C. Sufani, F. M. De Blasio, S. McDonald, and J. A. Rushby, "Validating the use of emotiv epoc in resting eeg coherence research," *Frontiers in Human Neuroscience*, no. 13.
[34] M. Egerstedt and X. Hu, "Formation constrained multi-agent control," *IEEE Trans. on Robotics and Automation*, vol. 17, no. 6, pp. 947–951, Dec 2001.
[35] M. K. Islam, A. Rastegarnia, and Z. Yang, "Methods for artifact detection and removal from scalp EEG: A review," *Neurophysiologie Clinique/Clinical Neurophysiology*, vol. 46, no. 4-5, pp. 287–305, Nov. 2016.
[36] A. Hyvarinen, "Fast and robust fixed-point algorithms for independent component analysis," *IEEE Trans. on Neural Networks*, vol. 10, no. 3, pp. 626–634, 1999.
[37] H. E. Wang, C. G. Benar, P. P. Quilichini, K. J. Friston, V. K. Jirsa, and C. Bernard, "A systematic framework for functional connectivity measures," *Frontiers in Neuroscience*, vol. 8, Dec. 2014.
[38] M. Vinck, M. van Wingerden, T. Womelsdorf, P. Fries, and C. M. Pennartz, "The pairwise phase consistency: A bias-free measure of rhythmic neuronal synchronization," *NeuroImage*, vol. 51, no. 1, pp. 112–122, May 2010.
[39] A. M. Bastos and J.-M. Schoffelen, "A Tutorial Review of Functional Connectivity Analysis Methods and Their Interpretational Pitfalls," *Frontiers in Systems Neuroscience*, vol. 9, Jan. 2016.
[40] M. Benedek, R. J. Schickel, E. Jauk, A. Fink, and A. C. Neubauer, "Alpha power increases in right parietal cortex reflects focused internal attention," *Neuropsychologia*, vol. 56, pp. 393–400, Apr. 2014.

LIST OF FIGURES

FOOTNOTES

Manuscript received ...

This work was supported by a Space Technology Research Institutes grant (grant number 80NSSC19K1052) from NASA's Space Technology Research Grants Program. *(Corresponding author: Zhaodan Kong)*

Gregory Bales *Student Member, IEEE*

Zhaodan Kong *Member, IEEE*

G. Bales and Z. Kong are with the Department of Mechanical and Aerospace Engineering, University of California, Davis, CA, 95616 (e-mail: zdkong@ucdavis.edu)