Differential Entropy Preserves Variational Information of Near-Infrared Spectroscopy Time Series Associated With Working Memory

Soheil Keshmiri[1]*, Hidenubo Sumioka[1], Ryuji Yamazaki[2] and Hiroshi Ishiguro[1,3]

[1] Hiroshi Ishiguro Laboratories, Advanced Telecommunications Research Institute International, Kyoto, Japan, [2] School of Social Sciences, Waseda University, Tokyo, Japan, [3] Graduate School of Engineering Science, Osaka University, Suita, Japan

Neuroscience research shows a growing interest in the application of Near-Infrared Spectroscopy (NIRS) in analysis and decoding of the brain activity of human subjects. Given the correlation that is observed between the Blood Oxygen Dependent Level (BOLD) responses that are exhibited by the time series data of functional Magnetic Resonance Imaging (fMRI) and the hemoglobin oxy/deoxy-genation that is captured by NIRS, linear models play a central role in these applications. This, in turn, results in adaptation of the feature extraction strategies that are well-suited for discretization of data that exhibit a high degree of linearity, namely, slope and the mean as well as their combination, to summarize the informational contents of the NIRS time series. In this article, we demonstrate that these features are inefficient in capturing the variational information of NIRS data, limiting the reliability and the adequacy of the conclusion on their results. Alternatively, we propose the linear estimate of differential entropy of these time series as a natural representation of such information. We provide evidence for our claim through comparative analysis of the application of these features on NIRS data pertinent to several working memory tasks as well as naturalistic conversational stimuli.

Keywords: near-infrared spectroscopy, differential entropy, NIRS time series feature extraction, brain activity decoding, working memory

1. INTRODUCTION

Recent years witness a growing interest in Near-Infrared Spectroscopy (Ferrari and Quaresima, 2012; Dix et al., 2013) as a promising tool for analysis of the brain activity of human subjects. Cui et al. (2010a) define NIRS as a technology for functional brain imaging based on hemodynamic signals from the cortex. The operational principle of NIRS devices is based on recording of the optical absorption of light over time to estimate the functionally evoked changes in cerebral oxy/deoxy-hemoglobin concentrations that result from local cerebral vascular and oxygen metabolic effects during the brain activity (Huppert et al., 2009). This technology is successfully applied in a variety of research areas, ranging from monitoring of the cerebral and myocardial oxygen sufficiency (Jobsis, 1977) and effect of aging on working memory (Vermeij et al., 2012) to decoding of vigilance (Bogler et al., 2014), complexity analysis of the neural activity of the children with attention-deficit/hyperactivity disorder (ADHD) (Gu et al., 2017), behavioral differences

between the genders (Baker et al., 2016), functioning of frontal cortex (Ozawa et al., 2014; Perlman et al., 2015), as well as determination of the level of difficulty of the mental tasks (Verner et al., 2013), to name a few.

Utilization of NIRS for monitoring the brain activity becomes more attractive, considering the non-invasive operational setup of NIRS-related devices. These devices are easy to use with portable, light-weighted headsets that are comparatively more immune to body movement (Dieler et al., 2012) while preserving unrestrictiveness and accessibility along with compact experimental setting (Moriai-Izawaa et al., 2012). These advantages provide a tremendous opportunity for their adaptation in naturalistic experiments through elimination of the requirement for confining the human subject to fMRI chamber (Naseer and Hong, 2015; Shin et al., 2016). It is worthy of note that although NIRS has a weaker signal-to-noise ratio (SNR) in comparison with fMRI (Gagon et al., 2011), research suggests that the hemodynamic responses that it captures in the form of blood oxy/deoxy-genation hemoglobin highly correlates with the BOLD in fMRI (Strangman et al., 2002; Steinbrink et al., 2006; Cui et al., 2011). Moreover, this correlation is unaffected by the nature of cognitive tasks (Strangman et al., 2002; Steinbrink et al., 2006; Cui et al., 2011). Such spatio-temporal similarities between BOLD and blood oxygenation (i.e., Δ oxy-Hb) are further supported by a number of research findings (Okamoto et al., 2004; Strangman et al., 2002; Toronov et al., 2007).

An important characteristic that is attributable to NIRS and fMRI time series is the underlying linear property of the hemodynamic responses that are measured by these devices (Dale and Buckner, 1997; Henson et al., 2001; Henson, 2003; Penny et al., 2005), particularly in prefrontal cortex (PFC) (Braver et al., 1997; Mitchell et al., 2008; Fishbum et al., 2014). Friston et al. (1994a) first investigate this linear property of hemodynamic responses, leading to formalization of General Linear Model (GLM) (Friston et al., 2011). Later, Schroeter et al. (2004) and Plichta et al. (2007) independently validate the applicability of GLM to NIRS time series. Along the same axis, Kamran and Hong (2014) argue that NIRS time series is a linear combination of various components, ranging from dynamical characteristics of the oxy/deoxy-hemoglobin (HbO/HbR) changes in a specific brain region to the baseline effect (Gusnard and Raichle, 2001), thereby further extending the results in (Friston et al., 1994a; Plichta et al., 2007; Friston et al., 2011). The overall acceptance of such an underlying linear dynamics is strengthened, considering the adaptation of GLM and its variation in fNIRS-fMRI comparative studies (e.g., Steinbrink et al., 2006; Sato et al., 2013, 2016) as well as its inclusion in major f/NIRS statistical analysis and modeling toolboxes (Koh et al., 2007; Huppert et al., 2009; Strangmann, 2009; Ye et al., 2009; Fekete et al., 2011). Tak and Ye (2013) present a comprehensive review of the statistical analysis and modeling of f/NIRS time series.

These results, in turn, help explain the choice of feature extraction methodologies that are well-suited for discretization of data with a high degree of linearity (i.e., linear changes in hemodynamics in response to a given stimulus), namely, the slope (Herff et al., 2014), the mean (Fazli et al., 2012;

Khan et al., 2014) as well as their combination[1] (Naseer and Hong, 2013; Shin et al., 2016) to summarize the information content of NIRS time series. This claim is further supported by Cui et al. (2010b) whose comparative analysis suggest that the slope (i.e., a linear correlate) of the NIRS data forms a significantly informative feature in contrast to various feature spaces. However, averaging-based feature extraction strategies are ill-suited for capturing the essential information pertinent to brain responses to stimuli (Spiers and Maguire, 2007; Erceg-Hurn and Mirosevich, 2008; Hasson and Honey, 2012; Wilcox, 2012, 2017; Wehbe et al., 2014; Liu et al., 2017; Rousselet et al., 2017). In fact, Ben-Yakov et al. (2012) show the shortcoming of the averaging-based methods for detecting responses that are highly context dependent, resulting in worsening the effective signal-to-noise ratio. Moreover, Haynes and Rees (2007) claim that averaging of the brain activity loses its functional state at any given point in time. Rousselet et al. (2017) present a comprehensive survey and analysis on this matter.

On the other hand, research suggests a direct correspondence between variational behavior of the brain activity and its information content (Miller, 2001; Friston, 2010; Bastos et al., 2018; Lundqvist et al., 2018; Wutz et al., 2018). For instance, Miller (2001) argues that there is a direct correspondence between the "amount of information" and variance since "anything that increases the variance also increases the amount of information" (Miller, 2001). Along the same direction, Eden and Kramer Fano (1947) note the utility of Fano factor[2] in characterization of neural spiking. Similarly, Cohen et al. (2017) consider the ability to identify meaningful variation in data pertinent to brain activation as an indicator of an effective analysis approach. Accordingly, Lundqvist et al. (2018) show that effect of working memory tasks is significantly captured by the variance of the information content of neural activation, thereby suggesting changes in variational information of spike rate to best represent the burst of brain activity in response to WM tasks. These results, collectively, suggest that the variation in time series of brain activity summarizes the cognitive load that is undertaken by the brain in processing a given stimulus.

These findings are in line with the concept of entropy in information theory (Cover and Thomas, 2006), as originally formulated by Shannon (1948). It comes as no surprise that entropy in its various formulations (Zanin et al., 2012; Gao et al., 2015; Xiong et al., 2017) is utilized exhaustively for analysis of the complexity and information content of biological signals (Costa et al., 2002; Lungarella and Sporns, 2006; Sengupta et al., 2013). In fact, its application expands over a wide spectrum, ranging from analysis of the complexity of heartbeats (Richman and Moorman, 2000; Bian et al., 2012) and state of anesthesia (Silva et al., 2010) to physiological complexity of aging and disease (Goldberger et al., 2002b), detection of epileptic seizure (Srinivasan and Eswaran,

---

[1] Let m_i and s_i be the mean and slope of the ith segment. Then their combination, as it is practiced in NIRS analyses, implies the vector $v = [m_i, s_i]$.

[2] Fano factor or coefficient of variation is a measure of dispersion of a probability distribution of a random process, X, that is expressed as $F_X = \frac{\sigma_X^2}{\mu_X}$ with σ_X^2 and μ_X representing the variance and mean of X, respectively.

2007) and investigation of vigilance and emotional states (Zhang and Lu, 2015). In addition, there exists a number of excellent and comprehensive entropy-based toolboxes for analysis of such time series signals (Ince et al., 2009; Lindner et al., 2011; Lizier, 2014).

In particular, linear estimate of differential entropy (Xiong et al., 2017) is used in study and analysis of a number of neurophysiological signals such as cardiovascular control mechanism (Porta et al., 2015) and cardiorespiratory dynamics (Faes et al., 2015). Shi et al. (2013) suggest the applicability of differential entropy (DE) i.e., the entropy of a continuous random variable (Cover and Thomas, 2006) in analysis of electroencephalography (EEG) time series. Furthermore, Keshmiri et al. (2017) show that DE significantly improves the classification accuracy of NIRS time series pertinent to cognitive load in prefrontal cortex during working memory task in comparison with a number of feature extraction and classification strategies. These results suggest the utility of DE to the solution concept of decoding of physiological and neurophysiological signals. However, lack of a theoretical evidence in its level of association for capturing the dynamical information of such time series data is apparent. In other words, there is a paucity of research in realization of the rationale underlying the shortcomings of averaging-based feature extraction strategies and, subsequently, the utility of adaptation of information-theoretic and variational-based approaches in calculating the summary statistics of brain activity.

In this article, we address these shortcomings through a systematic investigation of the mathematical foundation of the degree of correspondence between DE and these time series. In line with this perspective, we argue that averaging-based feature spaces are inefficient in representing the information content of the NIRS time series of the brain activity of human subjects. In doing so, we adapt the viewpoint of Miller (2001), thereby synonymously interpreting the amount of information conveyed by a NIRS time series as its variation, other than its overall trend or expectation. We further validate our claim by demonstrating the insensitivity of averaging-based feature spaces to such variational information, implicit in NIRS time series. Subsequently, we demonstrate that the linear estimate of differential entropy (Xiong et al., 2017) of these series help resolve this shortcoming.

Our contributions are twofold. Firstly, we present the mathematical bases for the shortcomings of the averaging-based feature spaces. Secondly, we prove that efficiency of the linear estimate of differential entropy of the time series of brain activity is due to its functional correspondence with the underlying spiking rate of neural activity. This is in line with the results in the literature, implying the correspondence between brain regional activation and the increase of the blood flow (Gusnard and Raichle, 2001). Additionally, we show the effectiveness of DE through examination of the feature extraction of the time series of the brain activity of human subjects, performing several working memory (WM) tasks. These WM tasks include Listening Span Test (LST), N-Back, Stroop, and Mental Arithmetic (MA). Moreover, we examine the utility of DE for capturing the information content of NIRS time series associated with more

naturalistic scenarios using data pertinent to conversational tasks.

Our results suggest the potential that utilization of the linear estimate of the differential entropy of NIRS time series data can provide to the solution concept of the analysis of the brain activity of human subjects in response to working memory tasks whose variational intensities are nontrivial.

2. MATERIALS AND METHODS

2.1. Investigation of the Feature Spaces

As we presented in section 1, the central role of linear models in NIRS-based applications result in adaptation of the feature extraction strategies that are well-suited for discretization of data with a high degree of linearity, namely, the slope, the mean, and their combination to summarize the informational contents of the NIRS time series. In this section, we demonstrate the shortcomings of such feature spaces. Subsequently, we propose the use of linear estimate of DE of these time series as a natural choice for extracting the information content of these series.

It is apparent that slope is optimum if the data of a given time series is collinear and monotonic in its order. This is a substantial limiting factor since most physical system, including the brain, have their limits defined by their power limitation which corresponds to a limit on the variance of their outcome (Stone, 2015). In fact, Miller (2001) considers the amount of information, conceptually, synonymous with variance. Moreover, Pearl (1988) associates the expected cost of the best estimate of a given random variable to its variance. Furthermore, its significance is evident in a tremendous amount of research, dedicated to identifying the meaningful variation in the time series of the brain activity (Cohen et al., 2017). Given these remarks, it is plausible to interpret the efficiency of an adapted feature as its ability to capture the amount of variation in the data that it is applied on. We formalize the shortcoming of slope as a feature extraction strategy in capturing such a variational information of NIRS time series through following Proposition.

Proposition 2.1. *Slope is inefficient in capturing the variational information of the given time series.*

Proof: Slope, $\mathbb{m}$, of a time series, X, is a measure of its linear trend with respect to the deviation of the observations from the fitted line that best minimizes these deviations. As such, the variance of the slope represents the estimate of the variational information of the observations, given this fitted line, i.e., Devore (2009)

$$
\begin{aligned}
s_{\mathbb{m}}^2 &= \frac{\sum_{i=1}^{N}(x_i - \bar{x})(t_i - \bar{t})}{\sum_{i=1}^{N}(t_i - \bar{t})^2} \\
&= \frac{\sum_{i=1}^{N}(x_i - \hat{x}_i)^2}{(N-1)\sum_{i=1}^{N}(t_i - \bar{t})^2}
\end{aligned}
\tag{1}
$$

where $s_{\mathbb{m}}^2$ is the variance of the slope of the sample time series data $x_i \in X$ with $t_i \in T$ being its corresponding independent variable (e.g., time stamps associated with the sample data and/or their respective indices). $\|X\| = N = \|T\|$, $\bar{x}$, and $\bar{t}$ are their cardinality

and sample means, respectively. Furthermore, $\frac{\sum_{i=1}^{N}(x_i - \hat{x}_i)^2}{N-1}$ is the estimate of the variation of the observations from the fitted line and $N - 1$ is the degree-of-freedom. $\hat{x}_i$ indicates the vertical projection of the ith observation from the fitted line (i.e., its deviation). It is apparent that:

$$\sum_{i=1}^{N}(x_i - \hat{x})^2 \leq \sum_{i=1}^{N}(x_i - \bar{x})^2 \qquad (2)$$

Furthermore, for a monotonically increasing independent variable (e.g., time and/or indices of data point in time series), we have:

$$(N-1)\sum_{i=1}^{N}(t_i - \bar{t})^2 \geq (N-1) \Rightarrow \frac{1}{(N-1)\sum_{i=1}^{N}(t_i - \bar{t}_i)^2} \leq \frac{1}{N-1} \qquad (3)$$

Given the Equations (2) and (3), we have:

$$\frac{\sum_{i=1}^{N}(x_i - \hat{x}_i)^2}{(N-1)\sum_{i=1}^{N}(t_i - \bar{t})^2} \leq \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2 \qquad (4)$$
$$\Rightarrow s_{\hat{m}}^2 \leq s_X^2$$

where s_X^2 is the sample variance.

∎

A customary practice in NIRS analysis is the transformation of data to ensure the reduction of the effect of the overall brain activity that is unrelated to the events of interests. These include subtraction/division of data with resting data (Fazli et al., 2012; Verner et al., 2013) to cancel the effects such as default mode (Gusnard and Raichle, 2001; Fransson, 2006; Fox and Raichle, 2007) of the brain, z-normalization (Bogler et al., 2014) i.e., $\frac{x - \mu_X}{\sigma_X}$, $\forall x \in X$, for NIRS time series X with mean and standard deviation μ_X and σ_X, and scaling within $[0 \ldots 1]$ interval (Hong et al., 2015) i.e., $\frac{x - X_{min}}{X_{max} - X_{min}}$, $\forall x \in X$, with X_{min} and X_{max} indicating the minimum and maximum values in X. In fact, these steps are necessary to ensure reduction of undesirable effect such as biasing influence of outliers and/or residual effect of any previous cognitive load due to some unrelated mental activity. In other words, these steps help preserve the information gain from appropriate part of signal, given the adapted cognitive task. We utilize the concept of mutual information (Cover and Thomas, 2006; Stone, 2015) and its invariance under such transformations (Kinney and Atwal, 2014) to provide further evidence on inefficiency of slope as a feature extraction strategy for NIRS time series in the following Proposition.

Proposition 2.2. *Information gain based on the slope of a NIRS time series X under z-normalization and scaling is unwarranted.*

Proof: Let X be the original NIRS time series data and y its transformation such that:

$$y = g(X) \qquad (5)$$

where $g(X)$ is of the forms $\frac{x - \mu_X}{\sigma_X}$ or $\frac{x - X_{min}}{X_{max} - X_{min}}$, $\forall x \in X$ where μ_X, σ_X, X_{min}, and X_{max} are mean, standard deviation, minimum, and maximum values of X, respectively. The entropy of y is (Stone, 2015):

$$H(y) = H(X) + E[\log|\frac{\partial g}{\partial X}|] \qquad (6)$$

The mutual information between X and y is:

$$I(X, y) = H(X) - H(X|y)$$
$$= H(y) - H(y|X) \qquad (7)$$
$$= H(X) + E[\log|\frac{\partial g}{\partial X}|] - H(y|X)$$

This implies that:

$$H(X|y) = H(y|X) - E[\log|\frac{\partial g}{\partial X}|]$$
$$\Rightarrow H(X|y) - H(y|X) = -E[\log|\frac{\partial g}{\partial X}|] \qquad (8)$$

Furthermore, $I(X, y)$ attains its maximum (i.e., preservation of the information of the original random variable, X) when $\frac{\partial I(X,y)}{\partial g(X)} = 0$ i.e., its slope with respect to the transformation g is zero. Moreover, $g : X \rightarrow y$ without any external effect (e.g., difference between the input and output signals due to the effect of the medium). Therefore, the amount of uncertainty in y, given X must not exceed that of X:

$$H(y|X) \leq H(X|y) \Rightarrow H(X|y) - H(y|X) \geq 0 \qquad (9)$$

Using Equations (8) and (9), we have:

$$-E[\log|\frac{\partial g}{\partial X}|] \geq 0 \Rightarrow 0 \leq |\frac{\partial g}{\partial X}| < 1 \qquad (10)$$

and $\frac{\partial g}{\partial X}$ i.e., the slope is $\frac{1}{\sigma_X}$ or $\frac{1}{X_{max} - X_{min}}$, given the z-normalization or the scaling transformations, respectively. Therefore, Equation (10) holds iff $0 \leq \sigma_X < 1$, $\sigma_X \neq 0$ and $-1 \leq X_{max} - X_{min} < 1$, $X_{max} - X_{min} \neq 0$. However, this requirement is unwarranted since $X_{max} - X_{min} \in \mathbb{R}$ and $\sigma_X \in \mathbb{R}$.

∎

Although averaging provides an effective tool for the analysis of certain aspects of the naturalistic stimuli (e.g., neural coupling between the speaker and the listener, Stephens et al., 2010), this feature extraction strategy is ill-suited for capturing the essential information pertinent to the responses to such stimuli in their broader sense (e.g., storytelling and conversation, Spiers and Maguire, 2007; Erceg-Hurn and Mirosevich, 2008; Hasson and Honey, 2012; Wilcox, 2012, 2017; Wehbe et al., 2014; Liu et al., 2017; Rousselet et al., 2017). In fact, Ben-Yakov et al. (2012) show the shortcoming of the averaging-based methods for detecting responses that are highly context dependent (e.g., listening to a story), resulting in worsening the effective signal-to-noise ratio. Moreover, Haynes and Rees (2007) claim that the averaging of the brain activity loses its functional state at any given point in time. We formalize these observations through following Proposition.

Proposition 2.3. *Averaging results in an information representation that tends to the overall expected value of the underlying distribution of a given time series.*

Proof: Let X represent a given time series. Furthermore, let $x_i \subseteq X$ be the ith segment of X with length k i.e., $\|x_i\| = k \leq \|X\|$, and $\bar{x}_i \to \mu_X$ as $k \to \|X\|$, i.e.,

$$\lim_{k \to \|X\|} P(|\bar{x}_i - \mu_X| > \epsilon) = 0 \tag{11}$$

where $\bar{x}_i$ and μ_X are the sample mean of x_i and the mean of the time series X. $P(.)$ gives the probability of the occurrence of its argument and $|.|$ returns the absolute difference between its two parameters. It is apparent that there exists $\frac{\|X\|}{k}$ such segments and:

$$\mu_X = \frac{1}{\frac{\|X\|}{k}} \sum_{i=1}^{\frac{\|X\|}{k}} \bar{x}_i \tag{12}$$

due to the law of large numbers. Therefore, these values reveal the general trend of their respective segments, thereby tending to the overall trend i.e., the expected value of X as $\|X\| \to \infty$:

$$\lim_{\|X\| \to \infty} P(\mu_X - \frac{1}{\frac{\|X\|}{k}} \sum_{i=1}^{\frac{\|X\|}{k}} \bar{x}_i > \epsilon) = 0 \tag{13}$$

∎

A common practice in the literature pertinent to analysis of the effect of the stimuli on the hemodynamic and/or the neural activity of the human subjects is the application of the sliding window on the given time series that is associated with these signals. In such a setting, the two consecutive segments that are extracted from the time series X share an overlapping subsegment of length $0 \leq n \leq \|X\| - 1$. Here, zero indicates that the consecutive segments do not overlap, whereas the upper limit means that the original series divided into two with only a single data point uncommon between these segments (i.e., the minimum requirement for the segmentation of $\|X\|$ into two distinct segments). An immediate implication of Proposition 2.3 is that the larger the overlap between the two consecutive segments is, the closer their respective expected values are. In other words, the overlap between the two segments results in the greater loss of the variational information. Furthermore, it indicates that the averaging becomes less effective if the events that are associated with a stimulus are subtle in nature (e.g., emotional cues reflected in a story) as opposed to a rather trigger-based event (e.g., appearance of a specific digit on a blank screen at a well-defined, predetermined point in time).

On the other hand, Eden and Kramer (2010) note the Fano factor (Fano, 1947) as an important characterization of neural spiking. In addition, they point out that such action potentials (i.e., neuronal spikes) often exhibit a high variability. What the utilization of slope/mean of a NIRS time series falls short on representing is indeed this variability. It is apparent that these action potentials are in fact the underlying dynamics of

the brain activity, thereby reflecting the changes in the local hemodynamic such as blood oxygen content and its subsequent level of consumption (Gusnard and Raichle, 2001). In other words, fluctuations that are observed in NIRS time series are in fact an indirect representation of the responses that are induced by these action potentials as their underlying dynamics of the brain activity. In the following Proposition, we demonstrate the ability of the linear estimate of differential entropy (DE) (Cover and Thomas, 2006; Stone, 2015) in explicitly capturing this variability.

Theorem 2.4. *The expected rate of change in the linear estimate of differential entropy of a time series X is inversely proportional to the Fano factor, F.*

Proof: Linear estimate of the differential entropy of a random variable X is (Cover and Thomas, 2006; Xiong et al., 2017):

$$H(X) = \frac{1}{2} \log_b(g(X)) \tag{14}$$

where,

$$
\begin{aligned}
g(X) &= 2\pi e \sigma_X^2 \\
&= 2\pi e \frac{\sum_{x \in X}(x^2 - 2x\mu_X + \mu_X^2)}{N} \\
&= 2\pi e \left(\frac{\sum_{x \in X} x^2}{N} - \frac{2\mu_X \sum_{x \in X} x}{N} + \frac{N\mu_X^2}{N} \right) \\
&= 2\pi e \left(\frac{\sum_{x \in X} x^2}{N} - \mu_X^2 \right)
\end{aligned} \tag{15}
$$

with $N = \|X\|$ is the cardinality of X. Using Equation (15), derivative of H with respect to μ_X is:

$$
\begin{aligned}
\frac{\partial H}{\partial \mu_X} &= \frac{1}{2} \frac{\frac{\partial g}{\partial \mu_X}}{\ln(b) g(X)} \\
&= \frac{1}{2} \frac{\frac{\partial}{\partial \mu_X}\left[(2\pi e(\frac{\sum_{x \in X} x^2}{N} - \mu_X^2))\right]}{\ln(b) 2\pi e \sigma_X^2} \\
&= \frac{-4\pi e \mu_X}{\ln(b) 4\pi e \sigma_X^2} \\
&= \frac{-\mu_X}{\ln(b)\sigma_X^2} = -\frac{\mu_X}{\sigma_X^2} = -F^{-1}, \; b = e
\end{aligned} \tag{16}
$$

∎

2.2. Tasks

We conducted two series of experiments, referred to as *Working Memory Experiment* (WME) and *Conversational Task Experiment (CTE)*, hereafter. We chose WME due to its utility in analytical studies of the cognitive loads on the working memory of human subjects. Moreover, we considered a number of working memory tasks with well-established differences in their imposed cognitive load on the WM (Cui et al., 2011; Verner et al., 2013; Fishbum et al., 2014; Herff et al., 2014), thereby allowing for investigation of the ability of DE in capturing the underlying dynamics and variation of the brain activity in

response to differing cognitive loads. On the other hand, we chose CTE to analyze the utility of the linear estimate of DE in response to more subtle variation in brain activity in response to naturalistic stimuli (e.g., conversation, listening to a story, etc.). Results of these analyses contribute in realization of the utility of linear estimate of DE in analytical as well as decoding domains of NIRS-based modeling of brain activity of human subjects. Description of WME and CTE are as follows.

1. WME: It consisted of four different working memory tasks, namely, Listening Span Test (LST) (Osaka et al., 2003), N-Back (B), Stroop (S), and Mental Arithmetic (M). Each of these working memory (WM) tasks consists of two subtasks. Their descriptions are as follows.

   a. LST: There were two subtasks, namely, L1 and L2, consisting of two and three sentences, respectively. These sentences were readout to the participants sequentially. Participants were instructed to judge the validity of each sentence once its reading was over [e.g., Sun sets in the west. (yes/no?)]. Once reading of sentences of a given subtask were complete, participants were required to recall the first word of each of the sentences. This resulted in two and three words recall in case of L1 and L2, respectively.
   b. B: It included a one-back (B1) and a two-back (B2) WM tasks. We used a recorded call-out of numerical sequences (0 through 9) in which participants were required to respond to sequential (i.e., B1) and every-other (i.e., B2) repetition of these digits via clicking the arrow keys on the computer keyboard.
   c. S: It contained two subtasks, involving two-color (i.e., S1) and three-color (i.e., S2) streams. Both of these subtasks consisted of a sequence of twenty words (i.e., name of a color such as "red," or "green") that were randomly matched/mismatched with their corresponding colors (e.g., word "red" was shown with its matching color, red, or a mismatching color such as blue). We used the color/word "red, blue" in S1 and "red, blue, green" in S2.
   d. M: It comprised of two subtasks, requiring the mental addition of a two-digit number with a single-digit number (M1) and two two-digit numbers (M2), respectively. There were four addition operations in each of these subtasks, resulting in eight arithmetic operations in total. In case of M2, half of these operations resulted in carryover.

Every subject participated in all of these four WM tasks. We acquired a 1-min-long resting data of the participants (with their eyes closed) prior to start of each subtask which was followed by its corresponding task. Furthermore, we randomized the ordering of these WM tasks while keeping the order of their corresponding subtasks intact for all participants. We used the PsychoPy (Peirce, 2003) in WME.

2. CTE: This paradigm comprised of 3-min-long conversation sessions in which we discussed four different topics (two easy and two difficult) with the participants (in Japanese). We communicated with our participants through minimalist anthropomorphic android, the Telenoid, to eliminate the potential effect of human characteristics such as gender and age. This resulted in four separate sessions, per participant. In every session, we began with acquiring a 1-min-long resting data, followed by its corresponding 3-min-long experimental session. We kept the content of conversations intact in all sessions. Every subject participated in all of these settings. However, we randomized the order of the easy/difficult among our participants. We provided our participants with a 1-min-long resting break (while staying at their seat with their eyes closed) prior to the commencement of each of these session. We maintained an approximately 1.2 m distance between the seat of the participant and the Telenoid. A male person conversed with our participants in all four conversational sessions.

2.3. Subjects
1. WME: Thirteen young adults (nine females and four males, $M = 21.87$, SD $= 2.61$) participated in these tasks.
2. CTE: Our participants included twenty two individuals (fourteen females and eight males, $M = 46.55$, SD $= 10.62$).

All participants were right-handed [confirmed using FLANDERS (Nicholls et al., 2013) handedness questionnaire], were free of neurological and psychiatric disorders, and had no history of hearing impairment. Prior to the data collection, we received approval (approval code: 16-601-1) from the ethical committee at the Advanced Telecommunications Research Institute International (ATR), Kyoto, Japan. All subjects gave written informed consent in accordance with the Declaration of Helsinki. Subjects were seated on an easy armchair in the sound-attenuated experimental room, with instructions to fully relax and their eyes closed while resting.

2.4. Data Acquisition
We used Near-Infrared Spectroscopy (Ferrari and Quaresima, 2012; Dix et al., 2013) to collect the brain activities in the frontal area of our participants. We acquired the NIRS time series of our participants using a wearable optical topography system "HOT-1000," developed by Hitachi High-Technologies Corporation (please refer to **Figure 1**). Participants wore this device on their forehead in our experiments. This device collects data through four channels (i.e., $Left_1$, $Left_3$, $Right_1$, and $Right_3$, as shown in **Figure 1**). Furthermore, it allows for recording of the measurement of brain activity through detection of total blood flow via emitting a wavelength laser light (810 nm) at 10 Hz sampling frequency. The postfix numerical values that are assigned to these channels specify their respective source-detector distances. In other words, $Left_1$ and $Right_1$ have a 1.0 cm and $Left_3$ and $Right_3$ have 3.0 cm source-detector distances, respectively. Findings in the literature on brain region activation during memory and language processing suggest a left-lateralized activation in both genders with higher specificity in females (Haut and Barch, 2006; Li et al., 2010). Therefore, we primarily utilize the NIRS time series that is acquired through the channel $Left_3$ of this device in the present study. It is worth noting that the source-detector distance of 3.0 cm is adequate for proper data acquisition of cortical brain activity using NIRS-based
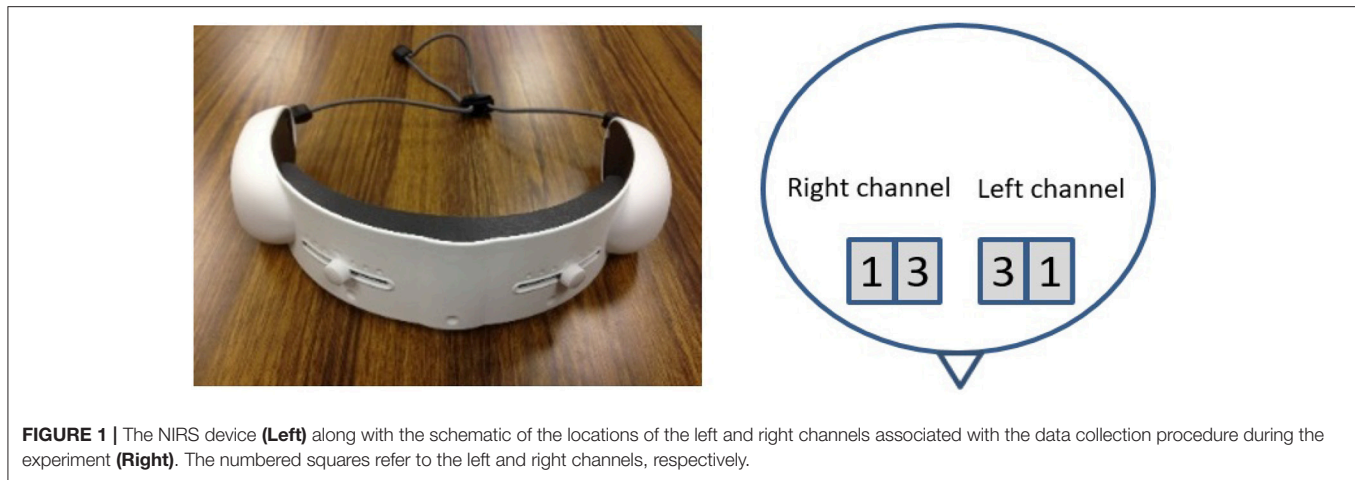
FIGURE 1 | The NIRS device **(Left)** along with the schematic of the locations of the left and right channels associated with the data collection procedure during the experiment **(Right)**. The numbered squares refer to the left and right channels, respectively.

devices (Ferrari and Quaresima, 2012; León-Carrión and León-Domínguez, 2012; Dix et al., 2013).

2.5. Data Preprocessing

First, we normalized the data corresponding to the selected NIRS channel via subtracting the mean of the 1 min resting period as a baseline from its data. Next, we applied a one-degree polynomial butter worth filter on this normalized data with 0.01 Hz and 0.6 Hz for low and high bandpass values, respectively. This was followed by its linear detrending.

2.6. Analyses

After data preprocessing step, we applied Wilcoxon signed-rank test on calculated features of time series data of our participants in both, WME and CTE experimental paradigms, to investigate the utility of different feature spaces in capturing potential differences in NIRS time series of brain activity. It is worth noting that we chose this non-parametric test to avoid any assumption on the underlying distribution of the data, as it is the case for two-sample t-test for example.

In case of WME, we first computed the mean, combined mean & slope, moving average, and DE features from the entire time series of corresponding subtasks, per participant. **Figure 2A**, illustrates this process. For combined mean & slope computation, we computed the mean and the slope of the entire time series, separately, and combined them into a single vector. We considered the slope of the fitted line to a given time series sequence for the latter. For moving average computation, we started with the first 5-s-long segment of a given NIRS time series and applied a 1-s moving window (i.e., 4 s of overlap between consecutive segments), calculating the mean of each of these segments. **Figure 2C**, shows this process. We used the mean of these averages (i.e., mean value of entries of V in **Figure 2C**) in our analysis. Furthermore, we used the normalized values of these features in comparative analysis of effectiveness of different feature spaces after application of normalization.

In case of CTE, we segmented each time series data into non-overlapping 5-s-long segments. For each segment, we extracted a feature using mean, mean & slope, moving average, and DE. **Figure 2B**, depicts this process. For combined mean & slope computation, we computed the mean and the slope (i.e., per segment), separately, and combined them into a single vector. We considered the slope of the fitted line to a given 5-s-long time series segment for the latter. Then, we computed the averaged mean & slope vector of these 5-s-long segments. We followed the same protocol as in WME for computing moving average (i.e., **Figure 2C**). We applied Wilcoxon signed-rank test on mean value (i.e., mean of entires of V in **Figure 2B** in case of mean, mean & slope, and DE and **Figure 2C** in case of moving average) of these extracted features.

For cluster analysis in **Figure 5**, we segmented each time series data pertinent to conversational task (i.e., CTE) into 20-s-long non-overlapping segments. For each segment, we computed a feature for every 5-s-long non-overlapping data sequence, resulting in a feature vector of length four in case of mean and DE (i.e., eight in combined mean & slope) for each 20-s worth of NIRS time series. **Figure 3A**, illustrates this process for the ith 20-s-long segment. On the other hand, we started with the first 5-s-long subsegment of each of these 20-s-long segments and applied 1-s moving window (i.e., 4 s of overlap between every consecutive subsegments) to calculate the feature vector based on moving average strategy. This resulted in a feature vector of length sixteen, per 20-s-long worth of data (**Figure 3B**). Next, we applied K-mean (Liao, 2005) algorithm (with two clusters) on these feature vectors (i.e., V in **Figure 2A** through **Figure 2C**).

3. RESULTS

In this section, We validate our mathematical proofs on effectiveness of the linear estimate of differential entropy (DE) of NIRS time series in contrast with the conventional averaging-based feature extraction strategies through statistical analysis (Wilcoxon test) of the brain activity of human subjects. Our analyses pertain to two different experiments, namely, Working Memory Tasks Experiment (WME) and Conversational Tasks
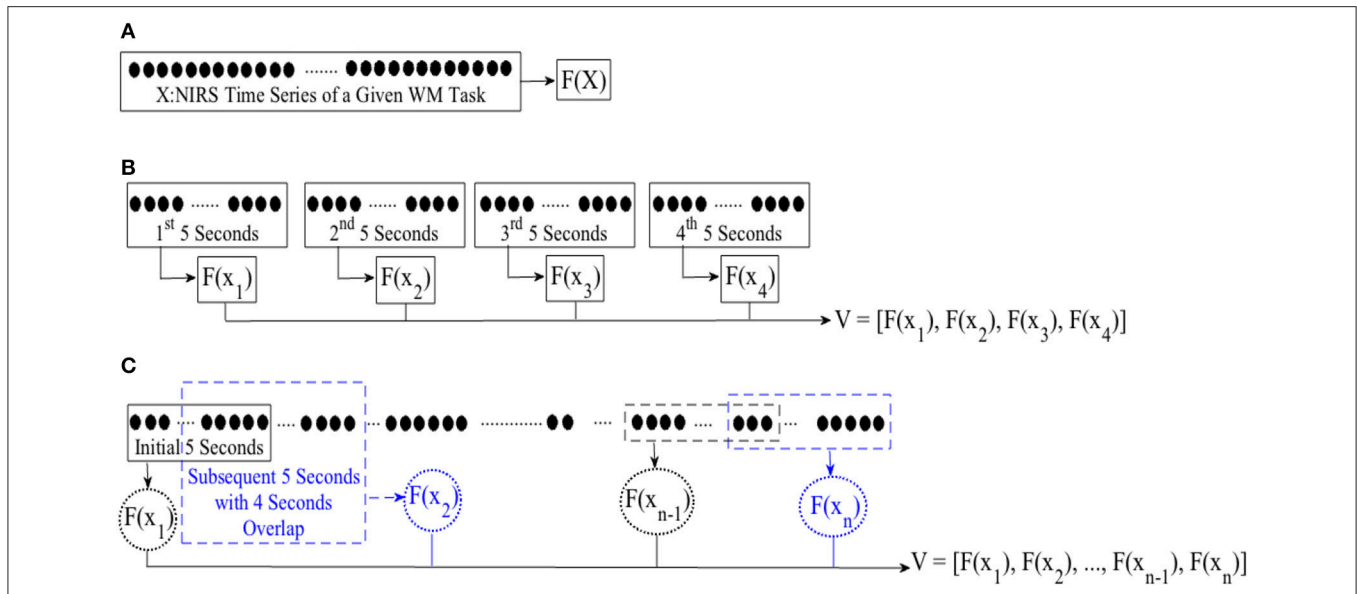
FIGURE 2 | Feature extraction from **(A)** entire NIRS time series **(B)** consecutive non-overlapping windows of length 5-s **(C)** first initial window of 5-s and following with subsequent 1-s moving windows (i.e., 4 s of overlap between consecutive sequences). $F(.)$ computes a desired feature from its argument, namely, differential entropy (DE), mean, slope, combination of mean & slope, or moving average. X is the entire time series of brain activity. $x_i \in X$ refers to ith 5-s-long segment of a given time series. V represents the feature vector, resulting from application of an adapted feature extraction strategy. For instance, $F(x_1)$ implies calculating mean of the 1st 5-s-long segment while adapting mean feature extraction strategy with V as its resulting mean-based feature vector. We used mean value of elements of V for statistical analyeses (i.e., both WME and CTE). On the other hand, we use V in its vector form during cluster analysis of brain activity pertinent to CTE.
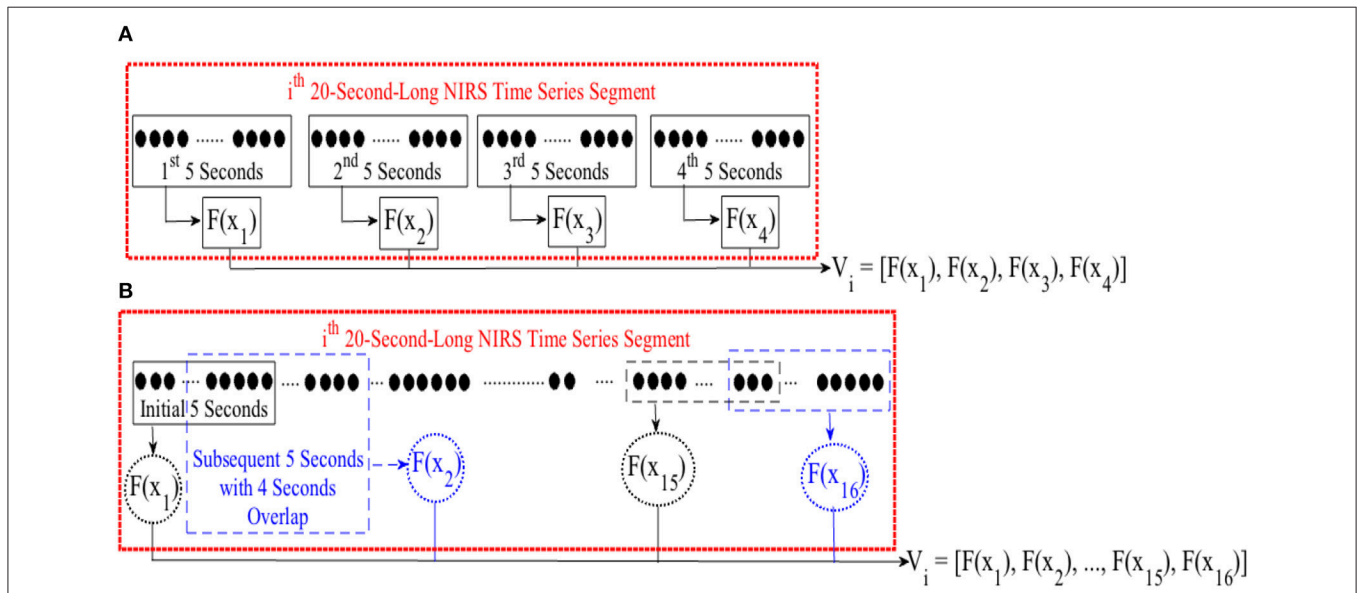
FIGURE 3 | (A) Feature calculation for every 5-s-long non-overlapping sequence of ith 20-s-long segment of NIRS time series of brain activity. This results in a feature vector V of length four in case of mean or DE feature spaces and length eight in case of combined mean & slope feature extraction strategy. **(B)** Extracting moving average feature vector. We started with first 5-s-long segment of NIRS time series data and applied a 1-s-long moving window (i.e., 4 s of overlap between consecutive sequences). For every 5-s-long segment, we computed the mean value. This resulted in a feature vector V of length sixteen. $F(.)$ computes a desired feature, namely, differential entropy (DE), mean, slope, combination of mean & slope, or moving average from its argument, X. $x_j \in X_i$ refers to jth 5-s-long segment of ith 20-s-long NIRS segment. V_i represents the feature vector of this ith 20-s-long segment. We applied K-mean (Liao, 2005) clustering (with two centroids) on these feature vectors during cluster analysis of CTE.

Experiment (CTE). Results in each of these sections are structured as follows.

1. WME: section 3.1.1 includes the analyses of Listening Span Test (LST), N-Back, Stroop, and Mental Arithmetic (MA) Working Memory (WM) tasks where each of these tasks comprised of two subtasks of different cognitive loads. In this subsection, we use features that are calculated based on actual value of NIRS time series data of the participants. On the other hand, we present results of this test on normalized NIRS time series data of the participants in section 3.1.2.

2. CTE: section 3.2 provide evidence on ability of linear estimate of DE in resolving the shortcoming of the averaging-based features in case of naturalistic stimuli (Spiers and Maguire, 2007; Erceg-Hurn and Mirosevich, 2008; Ben-Yakov et al., 2012; Hasson and Honey, 2012; Wilcox, 2012; Wehbe et al., 2014; Liu et al., 2017; Rousselet et al., 2017) via analyses of the brain activity of human subjects involved in four conversations with the content of whose varied in their level of difficulty.

3.1. WME
3.1.1. Original Feature Spaces

Wilcoxon test on application of mean as adapted feature space implied significant difference between two subtasks of Listening Span Test (LST) Working Memory (WM) tasks, i.e., L1 and L2 $[p < 0.001, T_{(24)} = -4.31]$, as well as subtasks S1 and S2 in Stroop WM $[p < 0.001, T_{(24)} = -6.65]$. However, it indicated non-significant with respect to Mental Arithmetic M1 and M2 $[p = 0.23, T_{(24)} = 1.20]$ as as well as N-Back B1 and B2 $[p = 0.47, T_{(24)} = 0.72]$ WM tasks.

Similarly, combination of mean & slope indicated significant difference between L1 and L2 $[p < 0.001, T_{(24)} = -4.31]$, as well as S1 and S2 $[p < 0.001, T_{(24)} = -6.65]$. However, it implied non-significant with respect to M1 and M2 $(p = 0.231124)$ as well as B1 and B2 $[p = 0.47, T_{(24)} = 0.72]$.

Although, application of moving average showed significant difference between L1 and L2 $[p < 0.001, T_{(24)} = -4.31]$ as well as S1 and S2 $[p < 0.001, T_{(24)} = -6.65]$, it implied non-significant

with regards to M1 and M2 $[p = 0.24, T_{(24)} = 1.17]$ as well as B1 and B2 $[p = 0.81, T_{(24)} = -0.23]$.

On the other hand, DE indicated significant differences between L1 and L2 $[p < 0.001, T_{(24)} = -4.31]$, S1 and S2 $[p < 0.01, T_{(24)} = -4.72]$, as well as M1 and M2 $[p = 0.03, T_{(24)} = -2.55]$ while suggesting non-significant difference between B1 and B2 $[p = 0.64, T_{(24)} = -0.47]$. **Table 1** provides summary statistics of these results.

3.1.2. Normalized Feature Spaces

Figure 4 shows discrimination of two subtasks of each of these WM tasks based on application of DE, mean, combined mean & slope, as well as moving average as feature extraction strategies on normalized NIRS time series data of the participants. In particular, **Figures 4A,B** show the ability of DE in capturing the significant different between the two subtasks in LST, i.e., L1 and L2 $[p < 0.05, T_{(24)} = -2.41, SD = 0.26]$ and Stroop, i.e., S1 and S2 $[p < 0.05, T_{(24)} = -1.99, SD = 0.22]$ where mean [LST: $p = 0.96, T_{(24)} = 0.05, SD = 0.27$, Stroop: $p = 0.15, T_{(24)} = -1.42, SD = 0.24$], mean & slope [LST: $p = 0.15, T_{(24)} = -1.43, SD = 0.27$, Stroop: $p = 0.15, T_{(24)} = -1.43, SD = 0.24$], as well as moving average [LST: $p = 1.0, T_{(24)} = 0.0, SD = 0.28$, Stroop: $p = 0.08, T_{(24)} = -1.65, SD = 0.23$] were unable to determine such differences. Most interesting is the result of utilization of these features in analysis of mental arithmetic (MA) where DE indicated an apparent significant between M1 and M2 $(p < 0.05, T_{(24)} = -2.02, SD = 0.28]$ while other features implied a tendency [Mean: $p = 0.15, T_{(24)} = 1.43, SD = 0.28$, Mean & Slope: $p = 0.15, T_{(24)} = 1.43, SD = 0.28$, Moving Average: $p = 0.14, T_{(24)} = 1.48, SD = 0.28$].
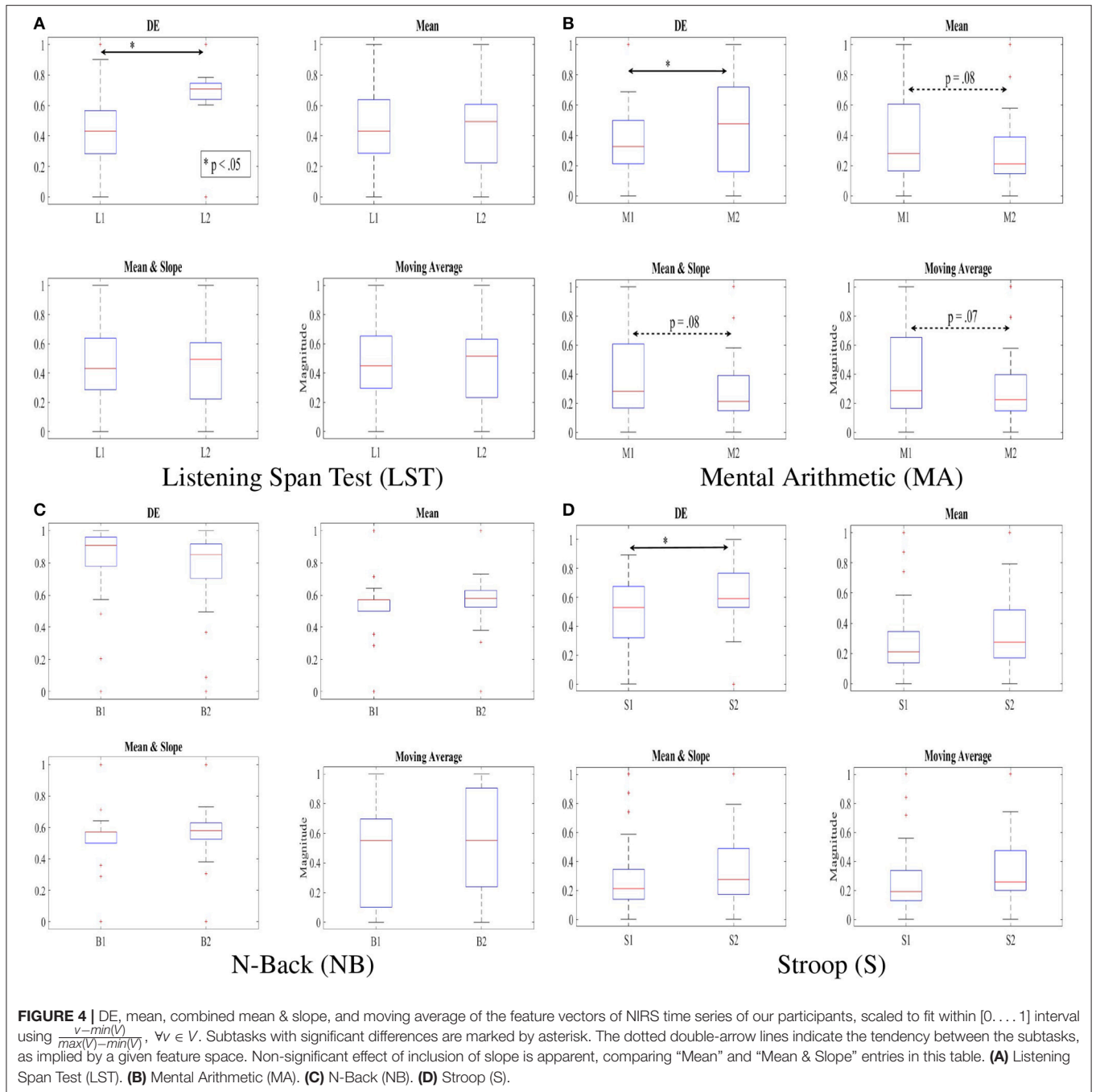
3.2. CTE

Wilcoxon test implied non-significant between NIRS time series of the participant with respect to the topic of conversation, i.e., easy and hard, using mean $[p = 0.37, T_{(746)} = 0.76]$, moving average $[p = 0.41, T_{(746)} = 0.83]$, as well as combined mean & slope $[p = 0.32, T_{(746)} = 0.69]$. However, it indicated significant based on application of DE on these time series $[p < 0.001, T_{(746)} = 4.00]$. **Table 2** provides the summary statistics of the mean- and DE-based features in conversational task experiment.

TABLE 1 | Working Memory Experiment (WME): Mean (M), Standard Deviation (SD), and Standard Error (SE) of Moving Average, DE, Mean, and combined Mean & Slope feature spaces with respect to N-Back (B1 and B2), Listening Span Test (L1 and L2), Stroop (S1 and S2), and Mental Arithmetic (M1 and M2) WM tasks.

WME	Moving Average			DE			Mean			Mean & Slope		
	M	*SD*	*SE*	*M*	*SD*	*SE*	*M*	*SD*	*SE*	*M*	*SD*	*SE*
B1	763.20	2.54	0.48	10.39	0.20	0.04	374.54	0.13	0.02	374.54	0.13	0.02
B2	762.81	3.71	0.70	10.39	0.17	0.03	376.85	48.57	0.09	374.52	0.09	0.02
M1	87.83	26.95	5.01	5.86	1.12	0.21	42.60	13.27	2.46	42.62	13.26	2.46
M2	80.44	29.06	5.40	6.68	1.03	0.19	39.10	14.25	2.65	39.12	14.25	2.64
S1	106.03	14.70	2.68	6.22	0.77	0.14	50.90	7.23	1.32	50.91	7.23	1.32
S2	212.98	27.80	5.08	7.51	1.24	0.23	105.82	14.12	2.58	105.82	14.12	2.58
L1	286.48	33.48	9.29	7.61	1.95	0.54	142.85	16.61	4.61	142.85	16.61	4.61
L2	771.07	0.69	27.93	10.47	0.31	0.09	376.85	48.57	13.47	376.85	48.57	13.47

These descriptive statistics are based on actual NIRS time series data of the participants (i.e., after preprocessing of the NIRS data and prior to application of normalization).

FIGURE 4 | DE, mean, combined mean & slope, and moving average of the feature vectors of NIRS time series of our participants, scaled to fit within [0. . . . 1] interval using $\frac{v - min(V)}{max(V) - min(V)}$, $\forall v \in V$. Subtasks with significant differences are marked by asterisk. The dotted double-arrow lines indicate the tendency between the subtasks, as implied by a given feature space. Non-significant effect of inclusion of slope is apparent, comparing "Mean" and "Mean & Slope" entries in this table. **(A)** Listening Span Test (LST). **(B)** Mental Arithmetic (MA). **(C)** N-Back (NB). **(D)** Stroop (S).

TABLE 2 | Conversational Tasks Experiment (CTE): Mean (M), Standard Deviation (SD), and Standard Error (SE) of Moving Average, DE, Mean, and combined Mean & Slope feature spaces with respect to easy and hard conversational topics during CTE.

CTE	Moving Average			DE			Mean			Mean & Slope		
	M	SD	*SE*	*M*	SD	*SE*	*M*	SD	*SE*	*M*	SD	*SE*
Easy	−20.31	179.05	9.46	1.73	0.84	0.04	−14.87	63.25	5.16	−14.83	61.20	7.16
Hard	−15.82	149.23	7.56	2.73	1.99	0.07	−8.62	76.61	9.26	−8.90	78.67	10.26

These descriptive statistics are based on actual NIRS time series data of the participants.

4. DISCUSSION

In this article, we argued that averaging-based feature extraction strategies that are inspired by high degree of linearity in NIRS time series of brain activity of human subjects results in suboptimal solution in capturing the variational information of these signals, thereby limiting the reliability of an adequate conclusion on their outcomes. Alternatively, we proposed the linear estimate of differential entropy of these time series as a natural representation of such information. We provided evidence for our claim through theoretical and empirical comparative analyses of the application of these features on NIRS data pertinent to a number of WM tasks with varying level of cognitive loads. Concretely, we demonstrated the utility of DE in contrast with mean, combination of mean & slope, and moving average feature spaces in analysis as well as differentiation of subtasks of a several WM tasks into their corresponding classes. These WM tasks included Listening Span Tests (L1 and L2), Stroop (S1 and S2), N-Back (B1 and B2), and Mental Arithmetic (M1 and M2). We further showed the confounding effect of these averaging-based feature extraction strategies via analysis of a naturalistic conversational tasks with differing level of difficulty in their respective topics, thereby providing evidence on inability of these features in representing the significance in responses of our participants to varying contextual complexity of conversational topics. Subsequently, we presented the sensitivity of DE in extracting this information. Moreover, we illustrated the substantial similarities between distribution of data based on mean and combination of mean & slope, thereby indicating the negligible contribution of the slope in representation of the information content of the brain activity of human subjects in response to WM as well as naturalistic conversational tasks with varying degree of difficulty in their topics.

Although we found similar indication of non-significant difference between N-back subtasks through application of DE as well as averaging-based feature spaces of the NIRS time series of the brain activity of the participants, we suggest that such a similarity is due to the significant resembling dynamics of these subtasks, and consequently, their equivalence in imposed cognitive loads on human subjects. This claim is due to comprehensive results in study and analysis of N-Back WM task (Owen et al., 2005; Fishbum et al., 2014). This observation is further supported by non-significant difference in induced level of complexity by B1 and B2 WM tasks on brain activity of our participants, as implied through result of multiscale entropy (MSE) analysis (Costa et al., 2002, 2005) of NIRS time series of their brain activity (please refer to Appendix for further details).

Fano factor (Fano, 1947) characterizes the neural spiking as a deviation of activation of neural population from their expected spiking rate. In other words, it signifies the expected brain activity in response to a given stimulus as per variation that is exhibited by neural population. In this regards, application of averaging-based feature spaces is equivalent to constraining such an activity within one standard deviation of its expected or average value (i.e., $\sigma^2 \leq 1$ in $F = \frac{\sigma^2}{\mu}$, Fano, 1947). It is

apparent that such a constraint is unwarranted, resulting in an information loss on functional state of neural population (Haynes and Rees, 2007). This interpretation finds evidence in recent findings that imply the effect of working memory tasks is significantly captured by the variance of the information content of neural activity (Lundqvist et al., 2018). In fact, change in brain activity of human subjects during WM tasks is quantified by its variational information (Miller, 2001) than its average activity or expectation. On the other hand, the mathematical bound between linear estimate of DE and Fano factor, as shown through Theorem 2.4, along with our empirical results on NIRS time series of brain activity during WM and conversational tasks suggest the effectiveness of DE in quantification of this variational information of brain activity of human subjects. However, further validation of these results on a larger sample size for drawing an informed conclusion on utility of DE in analysis of NIRS time series of brain activity of human subjects during WM tasks is necessary.

Another source of evidence on significance of the variational information of brain activity is due to the results of the analyses of physiological systems from perspective of their dynamical complexity. Research suggests that increase in complexity is an inherent attribute of healthy physiological systems (Lipsitz and Goldberger, 1992; Costa et al., 2002, 2005; Goldberger et al., 2002a,b; Takahashi et al., 2009). Additionally, it implies that such an increase in complexity strongly correlates with such cognitive functions as attention, memory, mental manipulation, verbal fluency, and language (Manor and Lipsitz, 2013; Yang et al., 2013). Interestingly, an increase in complexity implies a direct correspondence to variational information that is exhibited by such dynamical and physiological systems (Zhang, 1991; Fogedby, 1992; Gao et al., 2015). On the other hand, feature spaces such as event-specific mean activity and slope (i.e., a regression coefficient) follow the key assumption of linear model that the neuronal dynamics and their transients can be ignored (Friston et al., 1994a). As a result, they are suitable for scenarios in which neural responses are monotonic and, subsequently, fail if monotonicity in variational information is violated (Pouget et al., 2016).
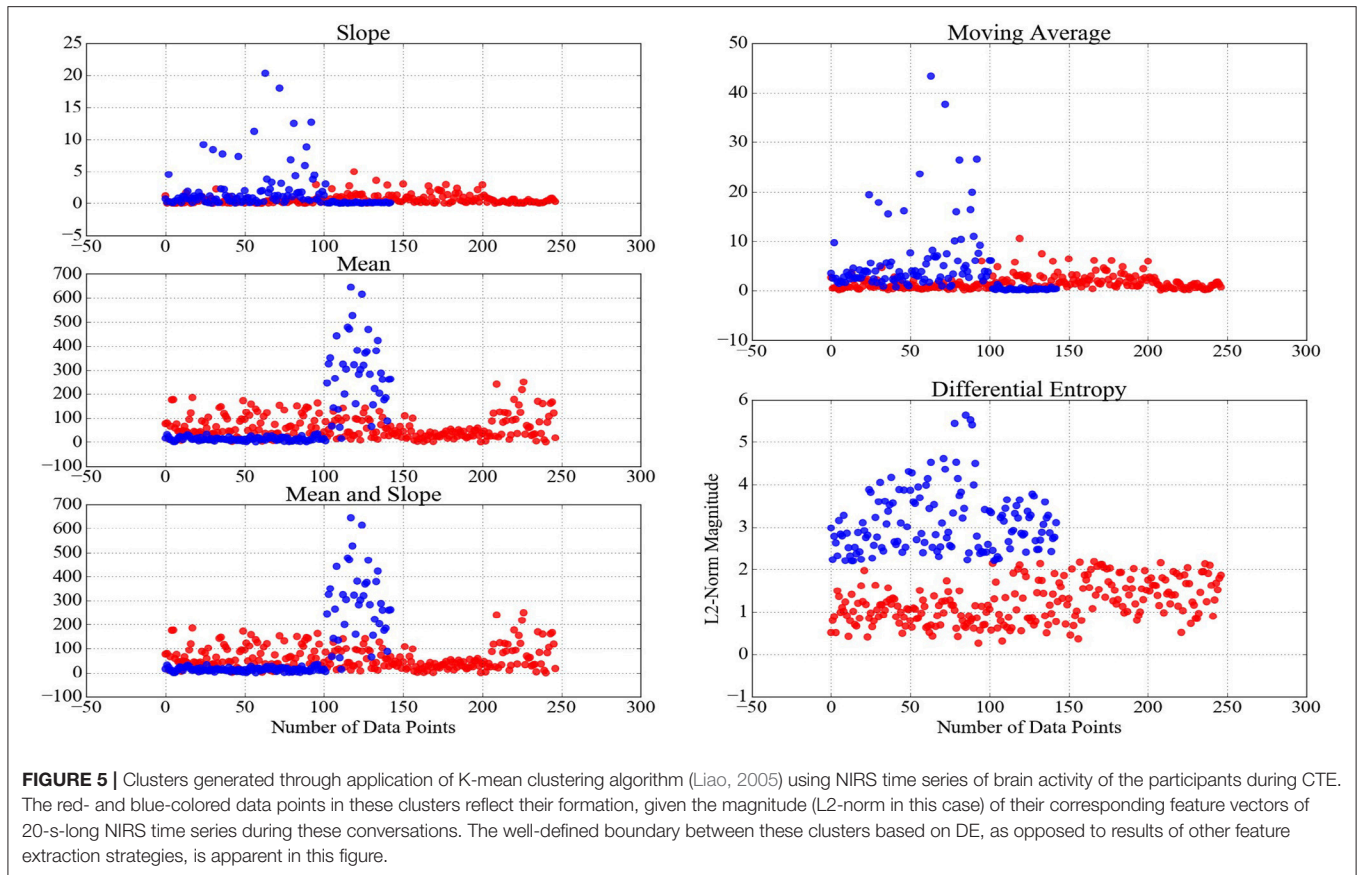
Research on working memory (WM) (Baddeley, 2003, 2012) suggests activation of prefrontal cortex (PFC) in a variety of tasks, ranging from mental tasks with high cognitive loads (Cohen et al., 1997; Tsujimoto et al., 2004; Owen et al., 2005) (i.e., change in activation in comparison with general/baseline activity, Gusnard and Raichle, 2001) to change in mental state (Ozawa et al., 2014; Sato et al., 2014) and social cognition such as emotional responses (Wolf et al., 2010) and story comprehension (Mar, 2011). An important implication of these results is the activation of some similar regions in PFC in response to these tasks. They are dorsolateral (MacPherson et al., 2002), ventrolateral, and medial PFC (Owen et al., 2005; Mar, 2011; Fishbum et al., 2014; Ozawa et al., 2014). Given the increased activity of PFC by mental tasks (Cohen et al., 1997; Tsujimoto et al., 2004; Owen et al., 2005; Ozawa et al., 2014; Sato et al., 2014), social cognition (Wolf et al., 2010) and story comprehension (Mar, 2011) along with these similar PFC activated regions in response to such tasks, it is plausible to attribute the ability of DE in preservation of

variational information of frontal brain activity in both, WM and conversational tasks, due to these common underlying activated regions (Owen et al., 2005; Mar, 2011; Fishbum et al., 2014; Ozawa et al., 2014). Considering the fact that DE is the average information content in a continuous time series data (Cover and Thomas, 2006; Avery, 2012; Stone, 2015), it is apparent that detected variation in such an information content summarizes the change of activity in these areas in response to stimuli (i.e., difficulty of mental tasks and topics of conversation in our case). Although configuration of recording channels of our NIRS device (please refer to **Figure 1**) to capture activity pertinent to PFC suggests further evidence in support of this observation, we are unable to confirm this presumption due to limited spatial resolution of our NIRS device. Therefore, future research to determine the credibility of this observation is required. It is also crucial to note that the ability of DE in preservation of variational information of tasks associated with WM does not warranty its utility as a universal NIRS feature, thereby necessitating further investigation to acquire insights on its performance with respect to NIRS time series of other mental states such as sleep and awake, relaxation and resting, and vigilance, to name a few.

An important implication of analytical studies of pattern of brain activity of human subjects during cognitive tasks is their integration in real-life applications (Mitchell et al., 2008; Baucom et al., 2012; Naseer and Hong, 2015; Shin et al., 2016; Horikawa and Kamitani, 2017). It is apparent that feature extraction strategies play a pivotal role in effectiveness of such decoding applications. This effect becomes more crucial, considering the requirement of the transformation of these time series (e.g., baseline- and z-transformation, scaling the feature vectors within $[0, \ldots, 1]$ interval, etc.) prior to their decoding, thereby preventing the detrimental effect of bias and variance. Therefore, it is desirable for an adapted feature extraction strategy to be able to maximize the preservation of the variational information of these time series data after the application of such transformations. For instance, Xiong et al. (2017) emphasize the substantial effect of normalization on detecting a fine-grained alterations of the dynamical transitions of physiological systems across time and multiple scales. In this respect, Proposition 2.2 implied that the adaptation of the slope of the NIRS time series of brain activity results in loss of such variational information. On the other hand, Keshmiri et al. (2017) suggest the ability of DE of NIRS time series in outperforming the averaging-based feature spaces for decoding of the brain activity during N-Back working memory task. Whereas their results provide evidence on utility of DE in supervised learning paradigms, it is of significant importance to examine its ability in preservation of the information content of the brain activity in tasks with the underlying cognitive load of whose exhibit higher degree of non-triviality due to higher subjectivity of responses of human subjects (e.g., naturalistic scenarios such as conversation and story comprehension). **Figure 5** shows two clusters that are generated using K-Mean clustering algorithm (Liao, 2005) on the feature vectors of the consecutive 20-s-long segments of NIRS time series of the brain activity of the participants during CTE. We used mean, slope, combination of mean and slope, moving average, and DE to generate these feature vectors to

determine the utility of different feature extraction strategies in clustering data pertinent to naturalistic stimuli (i.e., difficulty of conversational topics in this case). The y-axis of these subplots represent the magnitude of the L2-norm of vectors of each of these feature spaces. It is worthy of note that cardinality of these clusters do not reflect a one-to-one correspondence between the number of easy and difficult conversational topics in CTE, as defined by the overall presupposition of these topics in our experimental paradigm. However, we expect to observe such an uneven grouping behavior of the L2-norms of the time series of the brain activity of the participants, given the substantial differences of the mental responses of the individuals to subtleties and emotional cues in naturalistic scenarios (Haynes and Rees, 2007; Spiers and Maguire, 2007; Wolf et al., 2010; Ben-Yakov et al., 2012; Hasson and Honey, 2012; Wehbe et al., 2014; Liu et al., 2017). Considering the changes in the magnitude of these feature vectors, as summarized by their respective L2-norms, as well as the result of our analysis on the significance of the topics of conversation, it is plausible to interpret the grouping phenomenon that is exhibited by these clusters as evidence in effectiveness of DE in modeling the variational information of the brain responses of the participants during these naturalistic cognitive tasks. This interpretation is strengthened via comparison of DE-based clusters in contrast with clusters that are formed by mean, slope, combination of mean & slope, and moving average. In this figure, negligible informational contribution of the slope of these time series (as it was observed in case of WME) is evident. It is worthy of note that although our preliminary result on performance of DE in clustering the brain activity of human subjects during conversational tasks suggest its potential for fine-grained summarization of the information content of such time series data, future research is crucial to further investigate its property in related WM tasks, thereby determining its utility through more rigorous experimental paradigms with higher objective validation strategies (e.g., comparison of its classification results in contrast with self-assessed responses of the participants in response to their undergone cognitive load, their degree of interest in the topic and/or their perceived level of difficulty of the task, etc.).

Apart from the capability of DE in capturing the variational information that is implicit in the responses of the brain activity of human subjects to mental tasks with varying cognitive loads, DE presents a reliable tool for quantitative measurement of the amount of information in these activities. Concretely, we found that the difference in the amount of information (measured in bit i.e., base 2 logarithm) in B1 (Mean = 5.10, SD = 1.08) was significantly above one standard deviation from L1 (Mean = 3.77, $SD = 0.34$), M1 (Mean = 2.59, SD = 1.19), and S1 (Mean = 2.68, SD = 1.46), as observed in their z-score differences (0.20, 1.33, −0.8, −0.73). Similarly, M2 (Mean = 2.69, SD = 1.18) had an amount of information that was significantly one standard deviation below L2 (Mean = 5.15, SD = 1.07), B2 (Mean = 5.11, SD = 1.07), and S2 (Mean = 3.67, SD = 0.62), where their z-scores were 0.83, 0.90, −1.23, and −0.41, respectively. Furthermore, we found that such a difference was non-significant in case of conversational tasks (Seniors: Mean = 1.75, SD = 0.84,

FIGURE 5 | Clusters generated through application of K-mean clustering algorithm (Liao, 2005) using NIRS time series of brain activity of the participants during CTE. The red- and blue-colored data points in these clusters reflect their formation, given the magnitude (L2-norm in this case) of their corresponding feature vectors of 20-s-long NIRS time series during these conversations. The well-defined boundary between these clusters based on DE, as opposed to results of other feature extraction strategies, is apparent in this figure.

Min = 0.04, Max = 4.27 and Juniors: Mean = 2.48, SD: 1.80, Min: 0.12, Max: 7.58), with their z-scores within one standard deviation. These values of the information content of the time series data of our participants that are in well-agreement with the theoretical boundaries that are proposed for the capacity of the information processing of human subjects (Miller, 2001) provide further insight on dynamics of these WM tasks and their consequential effect on brain activity of human subjects. Additionally, it is plausible to interpret these differences in amount of information as a measure of their respective cognitive loads on the brain activity of the human subjects during these mental tasks.

Proposal of the linear estimate of differential entropy as a feature extraction strategy for NIRS time series implies two assumptions. Firstly, it assumes a linearity of the underlying dynamics of the time series data under investigation (Kaiser and Schreiber, 2002; Lizier, 2014). This is consistent with the widespread use of linear models on the basis of the linear contributions from different sources (Mitchell et al., 2008; Huppert et al., 2009; Friston et al., 2011). Secondly, it assumes the normality of the underlying distribution of such data. Although the validity of this assumption is unwarranted, a number of results on modeling fMRI time series help interpret its utility, given the observed correlation between NIRS hemodynamic blood oxy/deoxy-genation hemoglobin and BOLD (Strangman et al., 2002; Okamoto et al., 2004; Toronov et al., 2007; Cui

et al., 2011). In particular, Friston et al. (1994b) suggest that Poisson distribution well models the shape of the hemodynamic responses. Additionally, Boynton et al. (1996) show that gamma function is an effective tool in modeling hemodynamic responses in visual cortex. Furthermore, Aguirre et al. (1998) show that gamma function exhibits higher sensitivity in capturing the variability of these responses in an event-related sensorimotor task. The close relationship between the gamma and Poisson distribution is well understood. Concretely, for a random variable $X \sim \Gamma(\alpha, \beta)$, where α is an integer, we have $P(X \leq x) = P(Y \geq \alpha)$, $Y \sim Poisson(\frac{x}{\beta})$. Furthermore, if $Y \sim Poisson(\lambda)$, then time until its k arrivals is the gamma function $\Gamma(k, \frac{1}{\lambda})$, where k is the length of time until arrival of an event. Additionally, the correspondence between the Poisson and normal distribution is well realized. Precisely, an approximation of a Poisson distribution with parameter μ is the normal distribution $\mathcal{N}(\mu, \mu)$ (Rosner, 2016). In addition, the square root of a Poisson distributed random variable is approximately normally distributed (Johnson, 1993; Aguirre et al., 1998). Moreover, Poisson distribution forms a limiting case for a binomial distribution with a large number of observations and small probability of occurrence per observation while the z-score normalized observations of the latter correspond to the family of standard normal distribution. Furthermore, a binomially distributed random variable with parameters n and p is approximated by $\mathcal{N}(np, np(1 - p))$. These observations imply

the potential of the presence of an underlying normality of the hemodynamic responses in blood oxy/deoxy-genation of the NIRS time series, thereby presenting the potential of the linear estimate of differential entropy of these time series to address the shortcoming of the averaging-based strategies in WM and naturalsitic tasks (Spiers and Maguire, 2007; Ben-Yakov et al., 2012; Hasson and Honey, 2012; Wehbe et al., 2014; Liu et al., 2017).

5. CONCLUSION

In this article, we studied the shortcoming of averaging-based feature extraction strategies in capturing the information content of brain activity of human subjects, as represented by NIRS time series, during WM and conversational tasks. Furthermore, we demonstrated the efficiency of linear estimate of differential entropy (DE) in quantification of information content of such time series, thereby presenting its correspondence with the underlying spiking neural activity.

We validated our mathematical analyses through application of these features in analysis of a number of working memory (WM) tasks. Our results suggested that DE shows higher sensitivity to brain activity of human subjects in comparison with mean, slope, combination of mean & slope, as well as moving average feature spaces. In addition, we showed that DE has higher sensitivity with regards to information gain through comparative analysis of its results in contrast with averaging-based feature spaces after the application of normalization and scaling. It is worth noting that such steps as baseline correction, normalization, and scaling are of crucial importance since they help refine data, thereby reducing detrimental effect of undesirable variation such as effect of outliers and biasing.

We validated the sensitivity of DE in capturing the variational information of time series of brain activity of human subjects to more naturalistic scenarios through comparison of its results with respect to averaging-based feature spaces on data pertinent to conversational time series. Whereas mean, slope, combination of mean & slope, as well as moving average feature extraction strategies implied non-significant in brain activity of human subjects in response to topics of conversation (i.e., easy topic such as daily activities vs. difficult such as conversation on a controversial topic), DE indicated significant difference in

time series of brain activity in response to these conversational topics.

Although our results suggested the utility of DE in analysis of brain activity of human subjects pertinent to WM tasks with varying degree of cognitive loads, they do not imply its utility as a universal NIRS feature. For instance, it is necessary to examine its utility on NIRS time series of other mental states such as relaxation, resting, and vigilance, to name a few. Therefore, our results primarily represent the first step toward realization of the potential of this feature extraction strategy. Accordingly, future research that is devised with larger sample sizes along with more rigorous experimental settings and quantitative validation measures is necessary to derive an informed conclusion on its performance.

AUTHOR CONTRIBUTIONS

SK carried out the proofs and analyses. HS conceived the experiments and supervised their progress. RY conducted the conversational experiment. As the head of Hiroshi Ishiguro Laboratories (HIL), HI oversees the entire activity of all research teams and themes, ensuring the soundness of all proposals, quality of results, and their validity. SK and HS contributed equally in preparation of this manuscript.

REFERENCES

Aguirre, G., Zarahn, E., and D'Esposito, M. (1998). The variability of human, bold hemodynamic responses. *NeuroImage* 8, 360–369.

Avery, J. (2012). *Information Theory and Evolution, 2nd Edn*. Singapore: World Scientific Publishing Co. Pte. Ltd.

Baddeley, A. (2003). Working memory: looking back and looking forward. *Nat. Rev. Neurosci.* 4, 829–839. doi: 10.1038/nrn1201

Baddeley, A. (2012). Working memory: theories, models, and controversies. *Annu. Rev. Psychol.* 63, 1–29. doi: 10.1146/annurev-psych-120710-100422

Baker, J., Liu, N., Cui, X., Vrticka, P., Saggar, M., Hosseini, S. H., et al. (2016). Sex differences in neural and behavioral signatures of cooperation revealed by fnirs hyperscanning. *Sci. Rep.* 6:26492. doi: 10.1038/srep26492

Bastos, A., Loonis, R., Kornblith, S., Lundqvist, M., and Miller, E. (2018). Laminar recordings in frontal cortex suggest distinct layers for maintenance and control if working memory. *Proc. Natl. Acad. Sci. U.S.A* 115, 1117–1122. doi: 10.1073/pnas.1710323115

Baucom, L., Wedell, D., Wang, J., Biltzer, D., and Shinkareva, S. (2012). Decoding the neural representation of affective states. *NeuroImage* 59, 718–727. doi: 10.1016/j.neuroimage.2011.07.037

Ben-Yakov, A., Honey, C., Lerner, Y., and Hasson, U. (2012). Loss of reliable temporal structure in event-related averaging of naturalistic stimuli. *NeuroImage* 63, 501–506. doi: 10.1016/j.neuroimage.2012.07.008

Bian, C., Qin, C., Ma, Q., and Shen, Q. (2012). Modified permutation entropy analysis of heartbeat dynamics. *Phys. Rev. E* 85, 021906–1–021906–7. doi: 10.1103/PhysRevE.85.021906

Bogler, C., Mehnert, J., Steinbrink, J., and Haynes, J.-D. (2014). Decoding vigilance with NIRS. *PLoS ONE* 9:e101729. doi: 10.1371/journal.pone.0101729

Boynton, G., Engel, S., Glover, G., and Heeger, D. (1996). Linear systems analysis of functional magnetic resonance imaging in human v1. *J. Neurosci.* 16, 4207–4221. doi: 10.1523/JNEUROSCI.16-13-04207.1996

Braver, T., Cohen, J., Nystrom, L., Jonides, J., Smith, E., and Noll, D. (1997). A parametric study of prefrontal cortex involvement in human working memory. *NeuroImage* 5, 49–62. doi: 10.1006/nimg.1996.0247

Cohen, J., Dow, N., Engelhardt, B., Hasson, U., Li, K., Norman, Y. N., et al. (2017). Computational approaches to fMRI analysis. *Nat. Neurosci.* 20, 304–313. doi: 10.1038/nn.4499

Cohen, J., Peristein, W., Braver, T., Nystorm, L., Noll, D., Jonides, J., et al. (1997). Temporal dynamics of brain activation during a working memory task. *Lett. Nat.* 386, 604–607. doi: 10.1038/386604a0

Costa, M., Goldberger, A., and Peng, C. (2005). Multiscale entropy analysis of biological signals. *Phys. Rev. E* 71:021906. doi: 10.1103/PhysRevE.71.021906

Costa, M., Goldberger, A., and Peng, C.-K. (2002). Multiscale entropy analysis of complex physiologic time series. *Phys. Rev. Lett.* 89, 068102–1. doi: 10.1103/PhysRevLett.89.068102

Cover, T., and Thomas, J. (2006). *Elements of Information Theory, 2nd Edn.* New Jersey, NJ: John Wiley & Sons, Inc.

Cui, X., Bray, S., Bryant, D., Glover, G., and Reiss, A. (2011). A quantitative comparison of nirs and fMRI across multiple cognitive tasks. *NeuroImage* 54, 2808–2821. doi: 10.1016/j.neuroimage.2010.10.069

Cui, X., Bray, S., and Reiss, A. (2010a). Functional near infrared spectroscopy (NIRS) signal improvement based on negative correlation between oxygenated and deoxygenated hemoglobin dynamics. *NeuroImage* 49, 30–39. doi: 10.1016/j.neuroimage.2009.11.050

Cui, X., Bray, S., and Reiss, A. (2010b). Speeded near infrared spectroscopy (NIRS) response detection. *PLoS ONE* 5:e15474. doi: 10.1371/journal.pone.0015474

Dale, A., and Buckner, R. (1997). Selective averaging of rapidly presented individual trials using fmri. *Hum. Brain Mapp.* 5, 329–340.

Devore, J. (2009). *Probability and Statistics for Engineering and the Sciences.* Boston, MA: Brooks/Cole Cengage Learning.

Dieler, A., Tupa, S., and Fallgatter, A. (2012). Functional near-infrared spectroscopy for the assessment of speech related tasks. *Brain Lang.* 121, 90–109. doi: 10.1016/j.bandl.2011.03.005

Dix, L., Bel, F., BArets, W., and Lemmers, P. (2013). Comparing near-infrared spectroscopy devices and their sensors for monitoring regional cerebral oxygen saturation in the neonate. *Pediatr. Res.* 74, 557–563. doi: 10.1038/pr.2013.133

Eden, U., and Kramer, M. (2010). Drawing inferences from fano factor calculations. *J. Neurosci. Methods* 190, 149–152. doi: 10.1016/j.jneumeth.2010.04.012

Erceg-Hurn, D., and Mirosevich, V. (2008). Modern robust statistical methods: an easy way to maximize the accuracy and power of your research. *Am. Psychol.* 63, 591–601. doi: 10.1037/0003-066X.63.7.591

Faes, L., A.Porta, and Nollo, G. (2015). Information decomposition in bivariate systems: theory and application to cardiorespiratory dynamics. *Entropy* 17, 77–303. doi: 10.3390/e17010277

Fano, U. (1947). Ionization yield of radiations. II. The fluctuations of the number of ions. *Phys. Rev.* 72, 26–29. doi: 10.1103/PhysRev.72.26

Fazli, S., Mehnert, J., Steinbrink, J., Curio, G., Villringer, A., Muller, K.-R., et al. (2012). Enhanced performance by a hybrid NIRS-EEG brain computer interface. *NeuroImage* 59, 519–529. doi: 10.1016/j.neuroimage.2011.07.084

Fekete, T., Rubin, D., Carlson, J., and Mujica-Parodi, L. (2011). The NIRS analysis package: noise reduction and statistical inference. *PLoS ONE* 6:e24322. doi: 10.1371/journal.pone.0024322

Ferrari, M., and Quaresima, V. (2012). A brief review on the history of human functional near-infrared spectroscopy (fNIRS) development and fields of application. *NeuroImage* 63, 921–935. doi: 10.1016/j.neuroimage.2012.03.049

Fishbum, F., Norr, M., Medvedev, A., and Vaidya, C. (2014). Sensitivity of fNIRS to cognitive state and load. *Front. Hum. Neurosci.* 8:76. doi: 10.3389/fnhum.2014.00076

Fogedby, H. (1992). On the phase space approach to complexity. *J. Stat. Phys.* 69, 411–425. doi: 10.1007/BF01053799

Fox, M., and Raichle, M. (2007). Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nat. Rev. Neurosci.* 8, 700–711. doi: 10.1038/nrn2201

Fransson, P. (2006). How default is the default mode of brain function? Further evidence from intrinsic bold signal fluctuations. *Neuropsychologia* 44, 2836–2845. doi: 10.1016/j.neuropsychologia.2006.06.017

Friston, K. (2010). The free-energy principle: a unified brain theory? *Nat. Rev. Neurosci.* 11, 127–138. doi: 10.1038/nrn2787

Friston, K., Ashburner, J., Kiebel, S., Nichols, T., and Penny, W. (2011). *Statistical Parametric Mapping: The Analysis of Functional Brain Images.* London: Academic Press.

Friston, K., Holmes, A., Worsley, K., Poline, J., Frith, C., and Frackowiak, R. (1994a). Statistical parametric maps in functional imaging: a general linear approach. *Hum. Brain Mapp.* 2, 189–210. doi: 10.1002/hbm.460020402

Friston, K., Jezzard, P., and Turner, R. (1994b). Analysis of functional MRI time-series. *Hum. Brain Mapp.* 1, 153–171. doi: 10.1002/hbm.460010207

Gagon, L., Yucel, M., Dehaes, M., Cooper, R., Perdue, K., Selb, J., et al. (2011). Quantification of the cortical contribution to the NIRS signal over the motor cortex using concurrent NIRS-FMRI measurements. *NeuroImage* 59, 3933–3940. doi: 10.1016/j.neuroimage.2011.10.054

Gao, J., Hu, J., Liu, F., and Cao, Y. (2015). Multiscale entropy analysis of biological signals: a fundamental b-scaling law. *Front. Comput. Neurosci.* 9:64. doi: 10.3389/fncom.2015.00064

Goldberger, A., Amaral, L., Glass, L., Hausdorff, J., Ivanov, P., Mark, P., et al. (2000). Physiobank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals. *Circulation* 101, e215–e220. doi: 10.1161/01.CIR.101.23.e215

Goldberger, A., Amaral, L., Hausdorff, J., Ivanov, P., Peng, C., and Stanely, H. (2002a). Fractal dynamics in physiology: alterations with disease and aging. *Proc. Natl. Acad. Sci. U.S.A.* 99, 2466–2472. doi: 10.1073/pnas.012579499

Goldberger, A., Peng, C., and Lipsitz, L. (2002b). What is physiologic complexity and how does it change with ageing and disease? *Neurobiol. Aging* 23, 23–26. doi: 10.1016/S0197-4580(01)00266-4

Gu, Y., Miao, S., Han, J., Zeng, K., Ouyang, G., Yang, J., et al. (2017). Complexity analysis of fNIRS signals in adhd children during working memory task. *Sci. Rep.* 7:829. doi: 10.1038/s41598-017-00965-4

Gusnard, D., and Raichle, M. (2001). Searching for a baseline: functional imaging and the resting human brain. *Nat. Rev. Neurosci.* 2, 685–694. doi: 10.1038/35094500

Hasson, U., and Honey, C. (2012). Future trends in neuroimaging: neural processess as expressed within real-life contexts. *NeuroImage* 62, 1272–1278. doi: 10.1016/j.neuroimage.2012.02.004

Haut, K., and Barch, D. (2006). Sex influences on material-sensitive functional lateralization in working and episodic memory: men and women are not all that different. *NeuroImage* 32, 411–422. doi: 10.1016/j.neuroimage.2006.01.044

Haynes, J., and Rees, G. (2007). Decoding mental states from brain activity in humans. *Nat. Rev. Neurosci.* 7, 523–534. doi: 10.1038/nrn1931

Henson, R. (2003). *Human Brain Function, 2nd Edn.* Ch. 40, London: Academic Press.

Henson, R., Rugg, M., and Friston, K. (2001). The choice of basis functions in event-related fMRI. *NeuroImage* 13:149. doi: 10.1016/S1053-8119(01)91492-2

Herff, C., Heger, D., Fortmann, O., Hennrich, J., Putze, F., and Schultz, T. (2014). Mental workload during n-back task - quantified in the prefrontal cortex using fNIRS. *Front. Hum. Neurosci.* 7:935. doi: 10.3389/fnhum.2013.00935

Hong, K.-S., Naseer, N., and Kim, Y.-H. (2015). Classification of prefrontal and motor cortex signals for three-class fnirs-bci. *Neurosci. Lett.* 587, 87–92. doi: 10.1016/j.neulet.2014.12.029

Horikawa, T., and Kamitani, Y. (2017). Generic decoding of seen and imagined objects using hierarchical visual features. *Nat. Commun.* 8:15037. doi: 10.1038/ncomms15037

Huppert, T., Diamond, S., Franceschini, M., and Boas, D. (2009). Homer: a review of time-series analysis methods for near-infrared spectroscopy of the brain. *Appl. Opt.* 48, D280–D298. doi: 10.1364/AO.48.00D280

Ince, R., Peterson, R., Swan, D., and Panzeri, S. (2009). Python for information analysis theoretic analysis of neural data. *Front. Neuroinform.* 3:4. doi: 10.3389/neuro.11.004.2009

Jobsis, F. (1977). Noninvasive, infrared monitoring of cerebral and myocardial oxygen sufficiency and circulatory parameters. *Science* 198, 1264–1267. doi: 10.1126/science.929199

Johnson, N. (1993). *Univariate Discrete Distributions, 2nd Edn.* New York, NY: Wiley.

Kaiser, A., and Schreiber, T. (2002). Information transfer in continuous processes. *Physica D* 166, 43–62. doi: 10.1016/S0167-2789(02)00432-3

Kamran, M., and Hong, K. (2014). Reduction of physiological effects in fnirs waveforms for efficient brain-state decoding. *Neurosci. Lett.* 580, 130–136. doi: 10.1016/j.neulet.2014.07.058

Keshmiri, S., Sumioka, H., Yamazaki, R., and Ishiguro, H. (2017). A non-parametric approach to the overall estimate of cognitive load using nirs time series. *Front. Hum. Neurosci.* 11:15. doi: 10.3389/fnhum.2017.00015

Khan, M., Hong, M., and Hong, K.-S. (2014). Decoding of four movement directions using hybrid NIRS-EEG brain-computer interface. *Front. Hum. Neurosci.* 8:244. doi: 10.3389/fnhum.2014.00244

Kinney, J., and Atwal, G. (2014). Equitability, mutual information, and the maximal information coefficient. *Proc. Natl. Acad. Sci. U.S.A.* 11, 3354–3359. doi: 10.1073/pnas.1309933111

Koh, P., Glaser, D., Flandin, G., Kiebel, S., Butterworth, B., Maki, A., et al. (2007). Functional optical signal analysis: a software tool for near-infrared spectroscopy data processing incorporating statistical parametric mapping. *J. Biomed. Opt.* 12, 1–13. doi: 10.1117/1.2804092

León-Carrión, J., and León-Domínguez, U. (2012). Functional near-infrared spectroscopy (fNIRS): principles and neuroscientific applications. *Neuroimaging Methods, InTech*, 47–74. doi: 10.5772/23146

Li, T., Luo, Q., Gong, H., Roebers, C., and Rammsayer, T. (2010). Gender-specific hemodynamics in prefrontal cortex during a verbal working memory task by near-infrared spectroscopy. *Behav. Brain Res.* 29, 148–153. doi: 10.1016/j.bbr.2010.01.033

Liao, T. (2005). Clustering of time series data - a survey. *Patt. Recogn.* 39, 1857–1874. doi: 10.1016/j.patcog.2005.01.025

Lindner, M., Vicente, R., Priesemann, V., and Wibral, M. (2011). Trentool: a matlab open source toolbox to analyse information flow in time series data with transfer entropy. *BMC Neurosci.* 7:119. doi: 10.1186/1471-2202-12-119

Lipsitz, L., and Goldberger, A. (1992). Loss of complexity and aging. Potential applications of fractals and chaos theory to senescence. *JAMA* 267, 1806–1809. doi: 10.1001/jama.1992.03480130122036

Liu, Y., Piazza, E., Simony, E., Shewokis, P., Onaral, B., Hasson, U., et al. (2017). Measuring speaker-listener neural coupling with functional near infrared spectroscopy. *Sci. Rep.* 7, 1–13. doi: 10.1038/srep43293

Lizier, J. (2014). JIDT: an information-theoretic toolkit for studying the dynamics of complex systems. *Front. Robot. AI* 1:11. doi: 10.3389/frobt.2014.00011

Lundqvist, M., Herman, P., Warden, M., Brincat, S., and Miller, E. (2018). Gamma and beta bursts during working memory readout suggest roles in its volitional control. *Nat. Commun.* 9:394. doi: 10.1038/s41467-017-02791-8

Lungarella, M., and Sporns, O. (2006). Mapping information flow in sensorimotor networks. *PLoS Comput. Biol.* 2:e144. doi: 10.1371/journal.pcbi.0020144

MacPherson, S., Phillips, L., and Sala, S. D. (2002). Age, executive function and social decision making: a dorsolateral prefrontal theory of cognitive aging. *Psychol. Aging* 17, 598–609. doi: 10.1037/0882-7974.17.4.598

Manor, B., and Lipsitz, L. (2013). Physiologic complexity and aging: implications for physical function and rehabilitation. *Prog. Neuro-Psychopharmacol. Biol. Psychiatry* 45, 287–293. doi: 10.1016/j.pnpbp.2012.08.020

Mar, R. (2011). The neural bases of social cognition and story comprehension. *Annu. Rev. Psychol.* 62, 103–134. doi: 10.1146/annurev-psych-120709-145406

Miller, G. (2001). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Physiol. Rev.* 10, 343–352. doi: 10.1037/h0043158

Mitchell, T., Shinkareva, S., Carlson, A., Chang, K.-M., Malave, V., Mason, R., et al. (2008). Predicting human brain activity associated with the meanings of nouns. *Science* 320, 1191–1195. doi: 10.1126/science.1152876

Moriai-Izawaa, A., Danb, H., Dana, I., Sanoa, T., Ogurob, K., Yokotab, H., et al. (2012). Multichannel fnirs assessment of overt and covert confrontation naming. *Brain Lang.* 121, 185–193. doi: 10.1016/j.bandl.2012.02.001

Naseer, N., and Hong, K.-S. (2013). Classification of functional near-infrared spectroscopy signals corresponding to the right- and left-wrist motor imagery for development of a brain-computer interface. *Neurosci. Lett.* 553, 84–89. doi: 10.1016/j.neulet.2013.08.021

Naseer, N., and Hong, K.-S. (2015). fNIRS-based brain-computer interfaces: a review. *Front. Hum. Neurosci.* 9:3. doi: 10.3389/fnhum.2015.00003

Nicholls, M., Thomas, N., Loetscher, T., and Grimshaw, G. (2013). The flinders handedness survey (flanders): a brief measure of skilled hand preference. *Cortex* 49, 2914–2926. doi: 10.1016/j.cortex.2013.02.002

Okamoto, M., Dan, H., Shimizu, K., Takeo, K., Amita, T., Oda, I., et al. (2004). Multimodal assessment of cortical activation during apple peeling by nirs and fMRI. *NeuroImage* 21, 1275–1288. doi: 10.1016/j.neuroimage.2003.12.003

Osaka, M., Osaka, N., Kondo, H., Morishita, M., Fukuyama, H., Aso, T., et al. (2003). The neural basis of individual differences in working memory capacity: an fMRI study. *NeuroImage* 18, 789–797. doi: 10.1016/S1053-8119(02)00032-0

Owen, A., McMillan, K., Laird, A., and Bullmore, E. (2005). N-back working memory paradigm: a meta-analysis of normative functional neuroimaging studies. *Hum. Brain Mapp.* 59, 46–59. doi: 10.1002/hbm.20131

Ozawa, S., Matsuda, G., and Hiraki, K. (2014). Negative emotion modulates prefrontal cortex activity during a working memory task: a nirs study. *Front. Hum. Neurosci.* 8:46. doi: 10.3389/fnhum.2014.00046

Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference.* san francisco, CA: Morgan Kaufmann Publisher Inc.

Peirce, J. (2003). *Psychopy.* Available online at: http://www.psychopy.org

Penny, W., Ghahramani, Z., and Friston, K. (2005). Bilinear dynamical systems. *Philos. Trans. R. Soc. B Biol. Sci.* 360, 983–993. doi: 10.1098/rstb.2005.1642

Perlman, S., Huppert, T., and Luna, B. (2015). Functional near-infrared spectroscopy evidence for development of prefrontal engagement in working memory in early through middle childhood. *Cereb. Cortex* 26, 2790–2799. doi: 10.1093/cercor/bhv139

Plichta, M., Heinzel, S., Ehlis, A.-C., Pauli, P., and Fallgatter, A. (2007). Model-based analysis of rapid event-related functional near-infrared spectroscopy (NIRS) data: a parametric validation study. *NeuroImage* 35, 625–634. doi: 10.1016/j.neuroimage.2006.11.028

Porta, A., Faes, L., Marchi, A., Bari, V., Maria, B., Guzzetti, S., et al. (2015). Disentangling cardiovascular control mechanisms during head-down tilt via joint transfer entropy and self-entropy decompositions. *Front. Physiol.* 6:301. doi: 10.3389/fphys.2015.00301

Pouget, A., Drugowitsch, J., and Kepecs, A. (2016). Confidence and certainty: distinct probabilistic quantities for different goals. *Nat. Neurosci.* 19, 366–374. doi: 10.1038/nn.4240

Richman, J., and Moorman, J. (2000). Physiological time-series analysis using approximate entropy and sample entropy. *Am. J. Physiol. Heart Circul. Physiol.* 278, H2039–H2049. doi: 10.1152/ajpheart.2000.278.6.H2039

Rosner, B. (2016). *Fundamentals of Biostatistics, 8th Edn.* Boston, MA: Cengage Learning.

Rousselet, G., Pernet, C., and Wilcox, R. (2017). Beyond differences in means: robust graphical methods to compare two groups in neuroscience. *Eur. J. Neurosci.* 46, 1738–1748. doi: 10.1111/ejn.13610

Sato, H., Dresler, T., Haeussinger, F., Fallgatter, A., and Ehlis, A. (2014). Replication of the correlation between natural mood states and working memory-related prefrontal activity measured by near-infrared spectroscopy in a german sample. *Front. Hum. Neurosci.* 8:37. doi: 10.3389/fnhum.2014.00037

Sato, H., Yahata, N., Funane, T., Takizawa, R., Katura, T., Atsumori, H., et al. (2013). A nirs-fmri investigation of prefrontal cortex activity during a working memory task. *NeuroImage* 83, 158–173. doi: 10.1016/j.neuroimage.2013.06.043

Sato, T., Nambu, I., Takeda, K., Aihara, T., Yamashita, O., Isogaya, Y., et al. (2016). Reduction of global interference of scalp-hemodynamics in functional near-infrared spectroscopy using short distance probes. *NeuroImage* 141, 120–132. doi: 10.1016/j.neuroimage.2016.06.054

Schroeter, M., Bucheler, M., Muller, K., Uludag, K., Orbig, H., Lohmann, G., et al. (2004). Towards a standard analysis for functional near-infrared imaging. *NeuroImage* 21, 283–290. doi: 10.1016/j.neuroimage.2003.09.054

Sengupta, B., Stemmier, M., and Friston, K. (2013). Information and efficiency in the nervous system - a synthesis. *PLoS Comput. Biol.* 9:e1003157. doi: 10.1371/journal.pcbi.1003157

Shannon, C. (1948). A mathematical theory of communication. *Bell Syst. Tech. J.* 27, 379–423. doi: 10.1002/j.1538-7305.1948.tb01338.x

Shi, L., Jiao, Y., and Lu, B. (2013). "Differential entropy feature for eeg-based vigilance estimation," in *35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 6627–6630.

Shin, J., Muller, K.-R., and Hwang, H.-J. (2016). Near-infrared spectroscopy (NIRS)-based eyes-closed brain-computer interface (BCI) using

prefrontal cortex activation due to mental arithmetic. *Sci. Rep.* 6, 1–11. doi: 10.1038/srep36203

Silva, A., Cardoso-Cruz, H., Silva, F., Galhardo, V., and Antunes, L. (2010). Comparison of anesthetic depth indexes based on thalamocortical local field potentials in rats. *Anesthesiology* 112, 355–363. doi: 10.1097/ALN.0b013e3181ca3196

Spiers, H., and Maguire, E. (2007). Decoding human brain activity during real-world experiences. *Trends Cogn. Sci.* 11, 356–365. doi: 10.1016/j.tics.2007.06.002

Srinivasan, V., and Eswaran, C. (2007). Approximate entropy-based epileptic eeg detection using artificial neural networks. *IEEE Trans. Inform. Technol. Biomed.* 11, 288–295. doi: 10.1109/TITB.2006.884369

Steinbrink, J., Villringer, A., Kempf, F., Haux, D., Boden, S., and Obrig, H. (2006). Illuminating the bold signal: combined fMRI-FNIRS studies. *Magn. Reson. Imaging* 24, 495–505. doi: 10.1016/j.mri.2005.12.034

Stephens, G., Silbert, L., and Hasson, U. (2010). Speaker-listener neural coupling underlies successful communication. *Proc. Natl. Acad. Sci. U.S.A.* 107, 14425–14430. doi: 10.1073/pnas.1008662107

Stone, J. (2015). *Information Theory: A Tutorial Introduction.* Sheffield: Sebtel Press, Lightning Source UK Ltd.

Strangman, G., Culver, J., Thompson, J., and Boas, D. (2002). A quantitative comparison of simultaneous bold fmri and nirs recordings during functional brain activation. *NeuroImage* 17, 719–731. doi: 10.1006/nimg.2002.1227

Strangmann, G. (2009). Near-infrared neuroimaging with ninpy. *Front. Neuroinformatics* 3:12. doi: 10.3389/neuro.11.012.2009

Tak, S., and Ye, J. (2013). Statistical analysis of fnirs: a comprehensive review. *NeuroImage* 85, 72–91. doi: 10.1016/j.neuroimage.2013.06.016

Takahashi, T., Cho, R., Murata, T., Mizuno, T., Kikuchi, M., Mizukami, K., et al. (2009). Age-related variation in eeg complexity to photic stimulation: aa multiscale entropy analysis. *Clin. Neurophysiol.* 120, 476–483. doi: 10.1016/j.clinph.2008.12.043

Toronov, Y., Zhang, X., and Webb, A. (2007). A spatial and temporal comparison of hemodynamic signals measured using optical and functional magnetic resonance imaging during activation in the human primary visual cortex. *NeuroImage* 34, 1136–1148. doi: 10.1016/j.neuroimage.2006.08.048

Tsujimoto, S., Yamamoto, T., Kawaguchi, H., Koizumi, H., and Sawaguchi, T. (2004). Prefrontal cortical activation associated with working memory in adults and preschool children: an event-related optical topography study. *Cereb. Cortex* 14, 703–712. doi: 10.1093/cercor/bhh030

Vermeij, A., van Beek, A., Rikkert, M. O., Claassen, J., and Kessels, R. (2012). Effects of aging on cerebral oxygenation during working-memory performance: a functional near-infrared spectroscopy study. *PLoS ONE* 7:e46210. doi: 10.1371/journal.pone.0046210

Verner, M., Hermann, M., Torche, S., C.M.Roebers, and Rammsayer, T. (2013). Cortical oxygen consumption in mental arithmetic as a function of task difficulty: a near-infrared spectroscopy approach. *Front. Hum. Neurosci.* 7:217. doi: 10.3389/fnhum.2013.00217

Wehbe, L., Murphy, B., Talukdar, P., Fyshe, A., Ramdas, A., and Mitchell, T. (2014). Simultaneous uncovering the patterns of brain regions involved in different story reading subprocesses. *PLoS ONE* 9:e112575. doi: 10.1371/journal.pone.0112575

Wilcox, R. (2012). Comparing two independent groups via a quantile generalization of the wilcoxon-mann-whitney test. *J. Modern Appl. Stat. Methods* 11, 296–302. doi: 10.22237/jmasm/1351742460

Wilcox, R. (2017). *Introduction to Robust Estimation and Hypothesis Testing.* New York, NY: Academic Press.

Wolf, I., Dziobek, I., and Heekeren, H. (2010). Neural correlates of social cognition in naturalistic settings: a model-free analysis approach. *NeuroImage* 49, 894–904. doi: 10.1016/j.neuroimage.2009.08.060

Wutz, A., Loonis, R., Roy, J., Donoghue, J., and Miller, E. (2018). Different levels of category abstraction by different dynamics in different prefrontal areas. *Neuron* 97, 1–11. doi: 10.1016/j.neuron.2018.01.009

Xiong, W., Faes, L., and Ivanov, P. (2017). Entropy measures, entropy estimators, and their performance in quantifying complex dynamics: effects of artifacts, nonstationarity, and long-range correlations. *Phys. Rev. E* 95, 062114-1–062114-37. doi: 10.1103/PhysRevE.95.062114

Yang, A., Huang, C.-C., Yeh, H.-L., Liu, M.-E., Hong, C.-J., Tu, P.-C., et al. (2013). Complexity of spontaneous bold activity in default mode network is correlated with cognitive function in normal male elderly: a multiscale entropy analysis. *Neurobiol. Aging* 34, 428–438. doi: 10.1016/j.neurobiolaging.2012.05.004

Ye, J., Tak, S., Jang, K., Jung, J., and Jang, J. (2009). NIRS-SPM: statistical parametric mapping for near-infrared spectroscopy. *NeuorImage* 44, 428–447. doi: 10.1016/j.neuroimage.2008.08.036

Zanin, M., Zunino, L., Rosso, O., and Papo, D. (2012). Permutation entropy and its main biomedical and econophysics applications: a review. *Entropy* 14, 1553–1577. doi: 10.3390/e14081553

Zhang, W., and Lu, B. (2015). Investigating critical frequency bands and channels for eeg-based emotion recognition with deep neural networks. *IEEE Trans. Auton. Mental Dev.* 7, 162–175. doi: 10.1109/TAMD.2015.2431497

Zhang, Y.-C. (1991). Complexity and $\frac{1}{f}$ noise. a phase space approach. *J. Phys. I* 1, 971–977. doi: 10.1051/jp1:1991180

APPENDIX

Multiscale Entropy Analysis of the WME

In our analyses, we found similar indication of non-significant difference between N-back subtasks through application of DE as well as mean of the NIRS time series of the brain activity of the participants. We further suggested that such a similarity is due to the significant resembling dynamics of these subtasks, and consequently the equivalence in their imposed cognitive loads on human subjects, as opposed to insensitivity of DE. In addition, we provided support for our claim from the comprehensive results on study and analysis of this N-Back WM task (Owen et al., 2005; Fishbum et al., 2014). In this Appendix, we present additional evidence in support of our claim through complexity analysis of the brain activity of our participants during N-Back WM task.

Referring to **Figure A1**, we observed non-significant difference in the complexity of the brain activity of the participants in response to B1 and B2 tasks, as illustrated through analysis of their respective multiscale entropy (MSE) (Costa et al., 2002, 2005). Whereas we observed significant difference in the level of complexity of the brain activity of the participants during LST [$p < 0.001$, $T_{(38)} = 4.21$, SD = 0.86], mental arithmetic [$p < 0.03$, $T_{(38)} = 2.19$, SD = 0.87] and Stroop [$p < 0.03$, $T_{(38)} = -2.37$, SD = 1.0], it was non-significant in case N-back WM tasks [$p = 0.17$, $T_{(38)} = 1.41$, SD = 0.18].

FIGURE A1 | Grand averages of the mutli-scale entropy (MSE) values of the NIRS time series data of our participants. We use the pattern length $m = 2$, similarity criterion $r = 0.15$ and the range of 1 through 20 for the scale factor, as suggested by Goldberger et al. (2000).