Check for updates

Cognitive Outcome Prediction in Infants With Neonatal Hypoxic-Ischemic Encephalopathy Based on Functional Connectivity and Complexity of the Electroencephalography Signal

Noura Alotaibi[1,2]*, Dalal Bakheet[1,2], Daniel Konn[3], Brigitte Vollmer[4,5] and Koushik Maharatna[1]

[1] School of Electronics and Computer Science, University of Southampton, Southampton, United Kingdom, [2] Department of Computer Science and Artificial Intelligence, University of Jeddah, Jeddah, Saudi Arabia, [3] Clinical Neurophysiology, University Hospital Southampton, Southampton, United Kingdom, [4] Clinical Neurosciences, Clinical and Experimental Sciences, Faculty of Medicine, University of Southampton, Southampton, United Kingdom, [5] Paediatric Neurology, Southampton Children's Hospital, Southampton, United Kingdom

Impaired neurodevelopmental outcome, in particular cognitive impairment, after neonatal hypoxic-ischemic encephalopathy is a major concern for parents, clinicians, and society. This study aims to investigate the potential benefits of using advanced quantitative electroencephalography analysis (qEEG) for early prediction of cognitive outcomes, assessed here at 2 years of age. EEG data were recorded within the first week after birth from a cohort of twenty infants with neonatal hypoxic-ischemic encephalopathy (HIE). A proposed regression framework was based on two different sets of features, namely graph-theoretical features derived from the weighted phase-lag index (WPLI) and entropies metrics represented by sample entropy (SampEn), permutation entropy (PEn), and spectral entropy (SpEn). Both sets of features were calculated within the noise-assisted multivariate empirical mode decomposition (NA-MEMD) domain. Correlation analysis showed a significant association in the delta band between the proposed features, graph attributes (radius, transitivity, global efficiency, and characteristic path length) and entropy features (Pen and SpEn) from the neonatal EEG data and the cognitive development at age two years. These features were used to train and test the tree ensemble (boosted and bagged) regression models. The highest prediction performance was reached to 14.27 root mean square error (RMSE), 12.07 mean absolute error (MAE), and 0.45 R-squared using the entropy features with a boosted tree regression model. Thus, the results demonstrate that the proposed qEEG features show the state of brain function at an early stage; hence, they could serve as predictive biomarkers of later cognitive impairment, which could facilitate identifying those who might benefit from early targeted intervention.

Keywords: brain connectivity, cognitive scores, graph theory, electroencephalography (EEG), entropy analysis, hypoxic-ischemic encephalopathy (HIE), noise-assisted multivariate empirical mode decomposition (NA-MEMD)

INTRODUCTION

Hypoxic-ischemic encephalopathy (HIE) is one of the most severe birth complications that causes neonatal brain damage. The incidence of HIE is approximately 1–6 per 1000 live births (Byeon et al., 2015). Moderate to severe encephalopathy often leads to death, cerebral palsy, or severe neurodevelopmental impairment. Neurodevelopmental impairment (NDI) is a composite outcome that includes cognitive, behavioral, educational, and motor impairments. Cognitive deficit is considered one of the most expected outcomes associated with NDI, featured by slow information processing speed, deficits in working memory, attention, and executive function (Slaughter et al., 2016). This substantially impacts the affected individual and their families, including education, social participation, employment, and quality of life.

Early identification of infants at high-risk can help to provide targeted early intervention that aims to improve cognitive outcomes by taking advantage of the neuroplasticity of the developing brain in early infancy.

Recently, there has been increasing interest in exploring methods for assessing brain function in early infancy and using them as an aiding tool for the early prediction of cognitive impairments. Neuroimaging techniques have been used in several studies to identify infants at high-risk of cognitive impairment (Slaughter et al., 2016; Moeskops et al., 2017; He et al., 2018). Alongside neuroimaging methods, electroencephalography (EEG) is suggested to be the current gold standard technique for studying brain activity as it is relatively inexpensive, portable, non-invasive, user-friendly, and comparatively easy to use. Several studies have examined the feasibility of using EEG analysis to predict the cognitive outcome. Kong et al. (2018) conducted a systematic review highlighting the two basic approaches currently adopted for the early prediction of cognitive outcomes. One is the analysis of EEG features to identify the biomarkers that could help binary classify the subject as either cognitively impaired or normal. The second is the analysis of EEG characteristics to estimate the specific scores for a continuous cognitive measure that could predict cognitive performance. Compared to binary classification, prediction of cognitive development reflects the difference between individuals in terms of brain function and the levels of cognitive impairment, rather than determining the group membership as in the case of classification, which could be more challenging (Sui et al., 2020).

Limited previous studies have shown that early quantitative analysis of EEG can satisfactorily predict long-term cognitive outcome. Lloyd et al. (2021) employed serial, multichannel video EEG to predict outcome in preterm infants by finding the association between grading of EEG background activity– where EEG was recorded soon after birth and continued over the first 3 days–and developmental scores, at 2 years of age. Suppiej et al. (2017) compared spectral EEG values of infants born near term with infants born at extremely low gestational age, aiming to investigate whether spectral EEG features were related to neurological outcomes. The EEG data was recorded at 35 weeks post-conception, and the outcome was evaluated at 1 year of age with the Griffiths' scales.

Cainelli et al. (2021) carried out a longitudinal 6-year study to evaluate the feasibility of neonatal spectral EEG in predicting developmental delay in premature infants. The EEG data was recorded at 35 weeks post-conception, and the outcome was assessed at 6 years using Wechsler Pre-school and Primary Scale of Intelligence III. West et al. (2005) conducted regression-based analysis to predict outcomes at 18 months of 44 preterm infants using the quantitative measure of EEG continuity recording in the first 4 days after birth. Kühn-Popp et al. (2016) investigated the relationship between brain maturation processes and language skills (evaluated at 48 months) using EEG coherence measured at 14 months.

Although these attempts have paved the way to using EEG in early prediction of cognitive development, methodological limitations hinder further progress. For example, EEG grading systems are still subjective and dependent on interpretation by an expert.

On the other hand, coherence-based measures quantitatively estimate the linear similarity of relative amplitude and phase between signals in a specific frequency range. Although coherence analysis provides information on the degree of synchrony of brain activity at different locations for each frequency, it suffers from several limitations. First, it fails to capture the intrinsic non-linearity of brain activity, is unsuitable for tracking non-stationary dynamics as it partly depends on the amplitudes of the signals and is susceptible to the volume conduction issue (Sweeney-Reed and Nasuto, 2007). Moreover, coherence relies on both the amplitude and phase in its calculation, and there is increasing evidence that considering the synchronization of phase alone and separating it from amplitude information may allow capturing the synchrony of temporal information between signals. This temporal locking of phases between neural signals is considered essential for analyzing the dynamic neural assemblies underlying cognitive processing (Sweeney-Reed and Nasuto, 2007).

Spectral power can quantitatively capture EEG characteristics that could objectively predict the associated cognitive outcomes. However, the employed approaches are based on Fourier transform, which requires linearity and stationarity of the signals, which is not the case with EEG signals. Consequently, such spectral analysis methods may give misleading amplitude-frequency distribution for non-linear and non-stationary data. In addition, the spectral analysis methods used in the literature are based on *a priori* basis, often selected according to traditional frequency bands, which are inconsistent among studies. For example, alpha-band was chosen to be from 8 to 12 Hz in David et al. (2004), from 8 to 13 Hz in Breakspear et al. (2004), from 8 to 14 in Babiloni et al. (2006), or subdivided into 6–10 Hz and 10–14 Hz ranges in Stam et al. (2003). These small changes in the frequency ranges of interest may result in potentially informative brainwaves being missed, specifically in the case of infants, due to the well-known variability between them and the older individuals in the neural oscillations of interest (Saby and Marshall, 2012). Furthermore, using *a priori* basis is critical for both non-linear and non-stationary data, as one cannot expect a predetermined basis to fit all the non-linear and non-stationary dynamics (Huang et al., 1998).

Thus, further work is required to employ quantification and analysis methods that consider the non-linearity and non-stationary characteristics of EEG, intending to find objective and reliable biomarkers of the cognitive deficits.

This study, therefore, aims to investigate the effectiveness of non-linear quantitative EEG (qEEG) features within the regression-based framework for predicting the cognitive outcomes in term-born infants with neonatal HIE. Specifically, phase-based functional brain connectivity estimated by weighted phase-lag index (WPLI) with graph metrics and complexity analysis measured by sample entropy (SampEn), permutation entropy (PEn), and spectral entropy (SpEn) are the two classes of features used in this study. Both sets of features were previously validated on earlier prediction of CP in at high-risk infants with neonatal HIE (Bakheet et al., 2021). This study uses WPLI and entropies features to find the association between neonatal EEG and the individual cognitive performance (which were completed in a follow-up visit at 24 months of age). These features are chosen because both could capture the complex characteristics of the EEG spectra, particularly non-linear and non-stationary properties. While WPLI quantifies the phase synchronization between distinct brain areas over time, providing a global view of the whole-brain networks, the entropies measure the complexity of each EEG independently, providing an understanding of the dynamic process underlying specific brain area.

In order to estimate such features, it is usually necessary to decompose EEG signals into narrowband components. This step is required since the calculation of WPLI and entropies from a complex signal, as in EEG, composed of multiple frequency oscillators, does not reveal the underlying non-linear dynamics of the signals (Takahashi et al., 2010; Bruña et al., 2018). Thus, noise-assisted multivariate empirical mode decomposition (NA-MEMD) method is adopted to decompose the EEG signals into intrinsic components. One advantage of using NA-MEMD is that it is a fully adaptive method and does not require *a priori* selection of the filter cut-offs. Naturally, this is a valuable property because it can tackle the well-known frequency range variability according to the experimental condition. Another advantage of using NA-MEMD is it capable of dealing with the non-stationary property of EEG

signals by demonstrating the temporal evolution of different frequency components.

Correlation analysis is performed to ascertain the statistical significance of graph-theoretical parameters of WPLI and entropies features in finding the association with cognitive scores. Then, the significant features are used to train and test the tree ensemble regression models: boosting and bagging to evaluate their predictability in later cognitive development.

The remainder of the article is organized as follows: the materials and methods used in this study are described in section "Materials and Methods." Results are analyzed in section "Results" and discussed in detail in section "Discussion." Section "Conclusion" concludes the article.

MATERIALS AND METHODS

This section describes the methodology of the proposed analysis to predict cognitive development at 2 years of age. First, a description of the study population and recruitment process is given, followed by a description of the experimental setup. An overview of the overall process of the proposed analysis including the description of pre-processing techniques, the basic concept of NA-MEMD, WPLI-based functional brain connectivity analysis, graph theory, and complexity analysis, are also provided. Then, the procedure of how these methods is employed to extract the desired features will be demonstrated. Finally, the regression models used in this study are presented. A schematic outline of the proposed analysis is depicted in **Figure 1**.

Participants

Thirty term-born infants with HIE treated with hypothermia were recruited in this study. EEG data was recorded on the neonatal intensive care unit within the first week after birth. Routine clinical neurodevelopmental follow-up assessment was conducted at 24 months of age using the Bayley Scales of Infant and Toddler Development III (BSITD-III). Of the 30 infants, 20 completed the follow-up assessment. The BSITD-III consists of three scales; motor, language and cognitive, and for this study we used the composite scores from the cognitive scales. The BSITD-III cognitive scores from those twenty infants ranged

FIGURE 1 | Schematic outline of proposed analysis for predicting cognitive outcomes at 2 years of age.
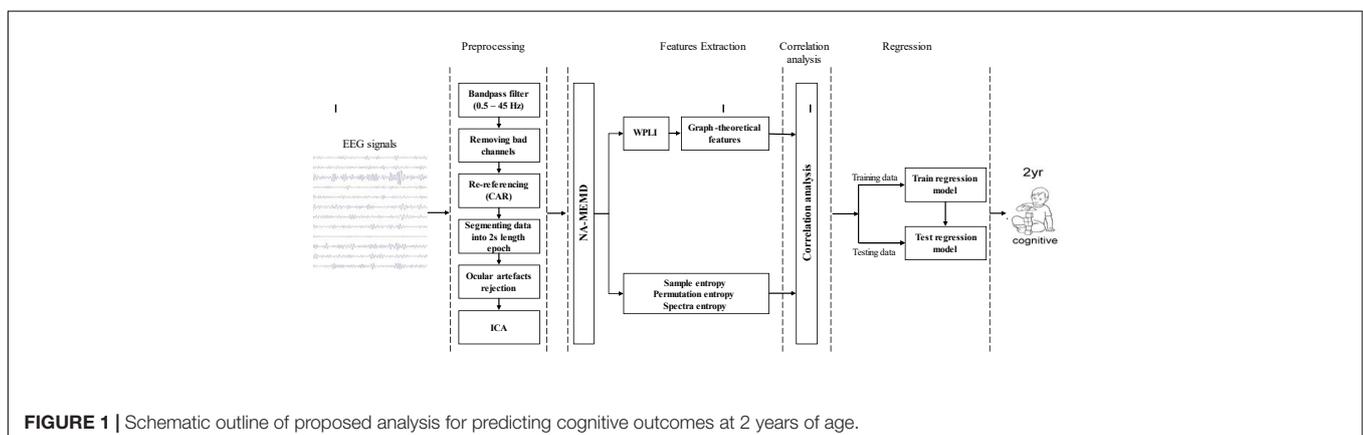
from 74 to 145. The BSITD-III mean of the normal population is 100, with a standard deviation (SD) of 1.5. Delay of cognitive development was defined as cognitive scores > 1.5 SD below the norm population mean. Ethical approval for secondary analysis of anonymized clinical data was obtained by the HRA and Health and Care Research Wales, HCRW (Reference ID 20/HRA/0260; IRAS project ID 278072; University Hospital Southampton R&D protocol number RHM CHI1047).

Data Acquisition

Electroencephalography data was recorded for 20 min during resting-state condition with eyes closed by either a Nihon Kohden (sampling frequency 1000 Hz, high-pass filter 0.016 Hz, the low-pass filter 300 Hz) and XLTEK (sampling frequency 512 Hz, high-pass filter 0.1 Hz, and the low-pass filter 70 Hz) clinical video-EEG system. Nineteen electrodes (C3, C4, CZ, F3, F4, F7, F8, FZ, FP1, FP2, O1, O2, P3, P4, PZ, T3, T4, T5, and T6) placed according to the 10–20 international system were used, as shown in **Figure 2**. Movement or electrode artifact affected the EEG in a substantial proportion of the cases. The first period in the EEG that was long enough without any clear significant artifact was always selected (the average length of the clips is approximately 2 min).

Data Pre-processing

The EEG data analysis was performed using MATLAB software package R2018a and EEGLAB toolbox. In order to improve the quality of the EEG signal, the remaining artifacts such as eye movement, muscle, heart activities, line noise, and signal discontinuity were eliminated using the following pre-processing techniques.

The EEG signals were initially recorded from 19 channels where the standard ground electrode was put close to Fz or Cz. After filtering the data, the EEG signals were automatically inspected to determine the set of bad channels. The procedure of picking the bad channels is based on the two criteria: first, the

flat channels, and second, the channels with a large amount of noise determined based on their standard deviation. Any channel marked as bad was eliminated and excluded from the further analysis. As a result of bad channels detection, the total number of removing channels were seven: C4, CZ, F4, F8, FP1, FP2, and Pz. The set of remaining channels used in calculating the common averaged reference (CAR) were 12: C3, F3, F7, Fz, O1, O2, P3, P4, T3, T4, T5, and T6. Thus, the brain topography used in this study included three brain regions: the central (C3, T3, and T4), anterior (F3, F7 and Fz), and posterior (O1, O2, P3, P4, T5, and T6) lobes.

A CAR was applied to re-reference the data to mitigate the confounding effect of the reference. According to a typical approach of EEG resting-state analysis, the EEG signal recorded from each channel was divided into epochs (windows), each of 2 s duration (Kułak and Sobaniec, 2003). This was chosen according to the natural properties of EEG, which are non-stationary in general and, however, it is quasi-stationary within only a short interval (Sakkalis, 2011). Thus, 2 s length appears to be the most appropriate length for capturing such EEG characteristics.

Epochs contaminated with ocular artifacts, particularly eye movement, were automatically identified using EEGLAB. A certain threshold of 55 µV was set and applied on each window; any epoch above the threshold was rejected and not used in further analysis (Apicella et al., 2013). Independent component analysis (ICA) was then applied, using the runICA algorithm implemented in EEGLAB, to remove the remaining artifacts from the signals, such as muscle artifacts and cardiac activity. Thus, the EEG signals from the 12 channels were separated into their 12 constituent independent components (ICs), as the general rule of ICA is to find the N ICs from the N linearly mixed signal (input channel data). These ICs are then projected back to the EEGs using the estimated separating matrix after eliminating the artifact-related ICs according to Chaumon et al. (2015). The algorithm of ICA was briefly described as follow:

The ICA decomposition finds an unmixing matrix (W) that decomposes the input channel data (x) into a sum of temporally independent and spatially fixed components, $u = Wx$. The rows of this output data matrix, u, called the component activations, are time courses of activation of the ICA components, and the columns of the inverse matrix W^{-1} give the projection strengths of the respective components onto the scalp sensors. The scalp topographies of the components provide information about the location of the sources (e.g., eye activity should project mainly to frontal sites, etc.). The classification of these components as artifact or EEG signal was performed using visual inspection based on the scalp topographies of the component. The artifact-free EEG data $x^{'}$ was fully reconstructed by multiplying of the inverse of W with $u^{'}$, where $u^{'}$ is the matrix of component activation, u, with rows representing artefactual components set to zero.

$$x^{'} \ = \ W^{-1}u^{'} \qquad (1)$$

The remaining epochs after the rejecting process were slightly varied between subjects. Since the infant who yields the lowest number of epochs upon rejection process gives 30 epochs, the first 30 epochs of each infant were considered. Thus, for each infant, a

FIGURE 2 | The 10–20 international system of EEG electrodes placement.
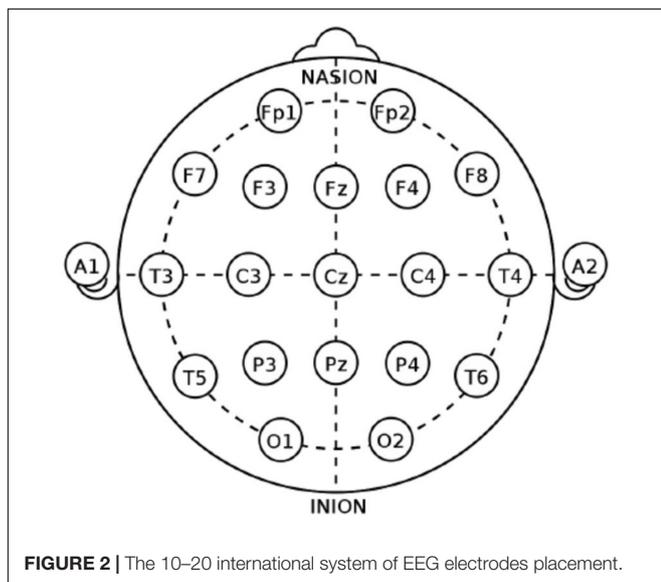
total of 12 channels, each with 30 artifact-free of 2 s EEG epochs, were used in the next stage of analysis.

Noise-Assisted Multivariate Empirical Mode Decomposition

Noise-assisted multivariate empirical mode decomposition is a data-driven time-frequency analysis capable to deal with non-stationary data (ur Rehman and Mandic, 2011). It was employed in this study to decompose the EEG signals into finite oscillation scales at the time domain. The proposed set of features was then calculated from each scale to characterize the time-frequency integration of information.

Noise-assisted multivariate empirical mode decomposition is a modified extension of empirical mode decomposition (EMD), figuring out its mode-alignment and mode-mixing problems (Huang et al., 1998). The mode-alignment refers to the issue of getting non-identical numbers of components when decomposing a multivariate signal, while the mode-mixing points out the situation of having different frequency ranges in a single scale. NA-MEMD solves the mode-mixing problem by adding a subspace containing multivariate independent white noise to the original multivariate signal, and then it processes the resulting composite signal using the multivariant empirical mode decomposition (MEMD) algorithm, which was proposed earlier to settle the mode-alignment (ur Rehman et al., 2010).

Unlike other traditional decomposing methods such as short-time Fourier (Gabor, 1946), wavelet transform (Mallat, 1989) and band-pass filters, EMD-based methods do not require a predefined basis of the signals. Instead, they decompose the time-series adaptively, through the Sifting process, from high to low-frequency components known as intrinsic mode functions (IMFs).

Among the available decomposition methods, the well-established wavelet analysis is known as one of the best non-stationary data analysis methods (Huang et al., 1998). However, the predefined basis of, for example, the Morlet wavelet (the most commonly used wavelet in general and in EEG analysis) leads to different issues (Sweeney-Reed and Nasuto, 2007). First, one cannot guarantee that the predetermined window size of the wavelet will coincide with a stationary period. Good localisation of the low-frequency oscillations needs a long-time window to identify them and thus a longer period of time for which signal should be stationary. On the other hand, selecting a small window may lead to missing potential biomarkers in the lower frequency ranges. Such a situation is known as the uncertainty principle, produced from the trade-off between frequency and time. Second, the prior selection of wavelet parameters cannot be expected to fit all the non-linear and non-stationary phenomena. Thus, it could induce spurious harmonic components to spectrally represent the signals, causing energy spreading and leading to faulty results.

Empirical mode decomposition-based methods satisfy the necessary conditions for the decomposition to represent a non-linear and non-stationary time series, particularly locality and adaptivity conditions. The locality is most crucial for non-stationarity, in which all events have to be identified by the time of their occurrences. Thus, both the amplitude (or energy) and the frequency are required to be functions of time (Huang et al., 1998). The adaptivity is important for both non-linear and non-stationary data in which the decomposition is adapted to the local variations of the data and hence can fully account for the underlying dynamics of the signals (Huang et al., 1998). Different studies proved that the local and adaptive nature of the decomposition using EMD-based methods is shown to improve time and frequency precision compared to the Morlet wavelet (Huang et al., 1998; Sweeney-Reed and Nasuto, 2007). A comparative summary of the Morlet wavelet and EMD-based methods is given in **Table 1**.

The procedure of the sifting process of the NA-MEMD method starts by considering a sequence of n-dimensional vectors $\{v(t)\}_{t=1}^{T} = \{v_1(t), v_2(t), v_3(t), \ldots, v_n(t)\}$ that represents a multivariate signal with n components (including the original signals and the added noise), and a set of direction vectors $X^{Q_k} = \{x_1^k, x_2^k, x_3^k, \ldots, x_n^k\}$ along the directions given by angles $Q^k = \{Q_1^k, Q_2^k, \ldots, Q_{(n-1)}^k\}$ on an $(n-1)$-sphere. Then the MEMD algorithm is applied as follows:

1) Choose a suitable set of points for sampling on a $(n - 1)$ sphere.
2) Calculate a projection, denoted by $\{P^{Q_k}(t)\}_{t=1}^{T}$, of the input signal $\{v(t)\}_{t=1}^{T}$ along the direction vector X^{Q_k}, for all k (the whole set of direction vectors), giving $\{P^{Q_k}(t)\}_{k=1}^{K}$ as the set of projections.
3) Find the time instants $t_j^{Q_k}$ corresponding to the maxima of the set of projected signals $\{P^{Q_k}(t)\}_{k=1}^{K}$.
4) Interpolate $[t_j^{Q_k}, v\left(t_j^{Q_k}\right)]$ to get the multivariate envelope curves $\{e^{Q_k}(t)\}_{k=1}^{K}$.
5) For a set of K direction vectors, the mean $m(t)$ of the envelope curves is calculated as $m(t) = \frac{1}{K} \sum_{k=1}^{K} e^{Q_k}(t)$.
6) Extract the detail $c_i(t)$ using $c_i(t) = v(t) - m(t)$ (i is an order of IMF). If the detail $c_i(t)$ satisfies the IMF conditions, apply the above procedure to $v(t) - c_i(t)$, otherwise apply it to $c_i(t)$.

The sifting process can be stopped when the detail $c_i(t)$ is monotonic and no more IMFs can be extracted from it.

Weighted Phase-Lag Index-Based Functional Brain Connectivity Analysis

Functional brain connectivity is an established technique for getting insight into the process of information propagation and

TABLE 1 | A comparative summary of the Morlet wavelet and EMD-based methods.

	Morlet wavelet	EMD-based methods
Basis	*a priori*	adaptive
Time-frequency precision	uncertainty	certainty
Non-linear	no	yes
Non-stationary	yes	yes
Theoretical base	yes	no (empirical)

relationship strength amongst the brain areas–the underpinning mechanism of the working principle of the brain. WPLI is the non-linear measure of functional brain connectivity that quantifies the phase synchronization between two signals (Vinck et al., 2011). It is an extended version of the phase-lag index (PLI) providing a better estimation of connectivity than PLI that is hindered by the discontinuity issue (Stam et al., 2007). This problem occurs in the case of small phase perturbation that could turn the phase lags into leads and vice versa (Vinck et al., 2011). WPLI was proposed to alleviate the effect of discontinuity of the connectivity index, volume conduction and other sorts of noise. It gives a reliable estimation of connectivity because it considers the magnitudes of the imaginary component of the cross-spectrum for weighting the phase differences between two sources of signals. Accordingly, the phase differences at high-risk for changing their true signs under the effect of small noise perturbations are assigned to small weight equivalent to the magnitude of the imaginary component. Subsequently, they would have a lower impact in quantifying connectivity.

Mathematically, WPLI can be defined as:

$$WPLI = \frac{|\langle |\Im(X)|\text{sign}(\Im(X))\rangle|}{\langle |\Im(X)|\rangle} \quad (2)$$

where $\Im(X)|$ is the imaginary component of the cross-spectrum X for two real-valued signals Z_1 and Z_2. The cross-spectrum X is computed as:

$$X = Z_1 * Z_2^* \quad (3)$$

where Z_2^* is a complex conjugate of Z_2. The value of WPLI ranges between 0 and 1, with 0 referring to no coupling between two signals, whereas one indicates that the two signals are perfectly coupled. WPLI quantified the functional brain connectivity between all twelve channels. Since a phase estimation would be better if it was extracted from a narrow frequency range in each source, in this study, the NA-MEMD method was adopted to decompose EEG signals into the intrinsic components. Then, these components were subjected to instantaneous phase estimation.

Fundamental Graph Theory

Graph theory is often applied to functional brain connectivity to describe the network architecture (Vecchio et al., 2017). In the graph theory concept, the brain can be represented as a network where the nodes correspond to distinct brain regions or EEG electrodes in EEG-based functional brain connectivity derivation, and the edges representing the functional connections between them. It can adequately characterize the network's topology and provide quantitative information about its properties. The graph-theoretical parameters measure these topological properties on both local and global scale. Local attributes identify the topological features of the single node, while the global metrics reveal the information flow over the whole network as well as any specialized local processing. There is increasing evidence that pathological conditions are viewed as a dynamic process affecting the entire brain.

On this basis, neuroimaging results have suggested that applying global attributes to quantify the global network

topological properties helps to reveal the disruptions in brain network behind such pathological conditions. Accordingly, in this study, the global parameters were chosen to investigate the whole topological properties of an infant's brain network. Identifying an aberrant in this network characteristics is expected to show the potential cognitive deficit emerging later during the child's lifespan. Mainly, transitivity, global efficiency, radius, diameter, and characteristic path length were used for this purpose. Global efficiency and characteristic path length are the measures of network integration referring to the ability of the network to transfer the information concurrently over the network. Radius and diameter provide insight into network eccentricity, while transitivity quantifies the ability of the network to localize information processing responsible for specific functions (Zhao et al., 2019). From the information processing perspective, networks possessing a high global efficiency and short characteristic path length have high efficiency in global information transfer and a high degree of network integration. On the other hand, networks with either a low radius or diameter have a high ability of information integration between brain areas. On another hand, the networks possessing high transitivity have a high local information transfer and these networks have a high tendency to specialize processing certain functions within a highly interconnected sub-network (Zhao et al., 2019). Even though the modularity measures the strength of the tendency of the network to divide into modules or groups, it is not considered in this study. This is because it can explain the capacity of a network of processing the local information rather than providing a view about the global information transfer within the network.

Those graph-theoretical parameters were computed using brain connectivity toolbox (BCT) in a MATLAB environment (Mika, 2010), and a brief description is illustrated in **Table 2**.

Complexity Analysis

Different studies in the literature report atypical EEG complexity associated with various neurodevelopmental disorders (Takahashi et al., 2010). Complexity analysis is utilized to provide a non-linear estimation of the dynamical brain activity. Entropy-based measures are commonly used to quantify time series complexity (Takahashi et al., 2010). Brief descriptions of the entropy measures employed in this study are given in the following.

Sample Entropy

Sample entropy, proposed by Richman and Moorman, provides an estimation of the irregularity or randomness of a time series (Richman and Moorman, 2000). It is a modified version of approximate entropy (ApEn) (Pincus, 1991), improving its immunity to the noise in the data and sensitivity to the signal length (Richman and Moorman, 2000). Both measures have been widely employed for the analysis of physiological data. SampEn measures the probability that two similar patterns for m point remain similar at the next $m+1$ point within a tolerance r. Thus,

TABLE 2 | List of graph parameters that used for characterizing a functional brain network.

Feature	Description
Transitivity	Reflecting connectivity of given region to its neighbors. The network with high transitivity implies it contains groups of regions that are densely connected internally.
Global efficiency	Representing the inverse of the shortest path between the regions. It measures the network efficiency in terms of how well the brain network integrated and how easily the information transfer between distinct brain regions.
Radius	Measuring the shape of network and it is the minimum of the network eccentricity.
Diameter	Measuring the shape of the network and it is the maximum of network eccentricity.
Characteristic path length	Representing the average distance between all pairs of brain regions in the network. It indicates how easily information transforms across the network.

for the time series $x(i)$ of length N, SampEn is given by:

$$SampEn\,(m, r, N) \;=\; -\ln[A^m(r)/B^m(r)], \quad (4)$$

where

$$A^m(r) \;=\; (N-m)^{-1}\sum_{i=1}^{N-m} C_i^{m+1}(r), \quad (5)$$

$$B^m(r) \;=\; (N-m)^{-1}\sum_{i=1}^{N-m} C_i^m(r), \quad (6)$$

$$C_i^m(r) \;=\; (N-m-1)^{-1}C_i,\; i \;=\; 1,\,2,\,..,N-m, \quad (7)$$

where m is the embedding dimension, $B^m(r)$ is the likelihood that $X_m(i)$ and $X_m(j)$ is matching for m points, while $A^m(r)$ is the likelihood that $X_m(i)$ and $X_m(j)$ will match for $m+1$ points. $C_i^m(r)$ is the probability of a vector $X_m(i)$ being similar to $X_m(j)$ within a tolerance r, C_i is the number that the distance two vectors $X(i)$ and $X(j)$ is smaller than r, and a vector $X_m(i)\,(1 \le i \le N-m+1)$ reconstituted of this series, and is given by: $X_m(i) \;=\; \{x(i),\,x(i+1),\,...,x(i+m-1)\}$. For optimal estimation of SampEn, some studies have recommended the embedding dimension $m = 2$ or 3, and the tolerance $r = 0.1–0.25$ of the standard deviation of the signal (Bruhn et al., 2000; Tudor et al., 2005). In this investigation, different parameter settings in these recommended ranges have been explored to check how robust the SampEn measures against these small changes in parameters.

Permutation Entropy

Bandt and Pompe (2002) developed PEn to determine the occurrence of ordinal patterns in time series data. The PEn is a robust and straightforward measure that quantifies the regularity of a time series by comparing neighboring values to estimate the intrinsic structures in EEG data (Bandt and Pompe, 2002). Thus, for the time series $x(i)$ of length N, the normalized PEn is given by:

$$PEn \;=\; -\frac{\sum_{i=1}^{n} p_i log p_i}{ln(n)} \quad (8)$$

where n is the order pattern, and p_i is the probability of the ith permutation occurring. The smaller the value of PEn, the more regular the time series is.

The appropriate selection of embedding parameters, including dimension m and time delay L, is necessary for proper PEn estimation. For this purpose, Olofsen et al. (2008) suggested the

values of $m = 3$, and $L = 1–2$. In this exploration, the PEn was computed using these recommended values to investigate whether the small changes in these parameters could affect the PEn estimation.

Spectral Entropy

Spectral entropy is another common EEG complexity measure that computes the randomness of the power spectrum of a signal (Li et al., 2008). Thus, unlike SampEn and PEn, SpEn measures the signal irregularity in the frequency domain. For this end, SpEn applies the Shannon entropy concept to the normalized power spectral density (PSD) of the signal such that,

$$SpEn \;=\; -\sum_{i=1}^{N} p_i \log p_i, \quad (9)$$

where p_i is the probability distribution of PSD at each frequency points and N is the total frequency points. SpEn was calculated using the frequency range 0.5–45 Hz. The frequency range was selected to focus on the range of interest with respect to traditional brain waves.

Spectral entropy is an efficient method to reflect the degree of skewness in the frequency distribution. A high value of SpEn indicates a flat, uniform spectrum with a broad spectral content, and a low value of SpEn describes a spectrum with all the power condensed into a single frequency point (Cui et al., 2015).

Features Extraction Procedure

The features extraction stage in the proposed analysis was divided into two parts. The first phase focused on decomposing EEG signals into their intrinsic components by using NA-MEMD. The second part involved extracting the two fundamentally different classes of features, namely graph-theoretical attributes and entropies features.

Step 1: Noise-Assisted Multivariate Empirical Mode Decomposition Analysis

1. A multivariate signal was constructed by combining the data points from all infants for each channel separately. The idea of combining signals from different sources decomposing them using NA-MEMD to acquire aligned IMFs was previously conducted in the literature (Zahra et al., 2010). This yielded twelve different matrices (i.e., one matrix per channel); each of them has the dimensionality of

$N_s \times N_t \times N_e$, where N_s denotes the number of subjects (which is 20), N_t indicates the number of temporal samples (which is 1024), and N_e is the number of epochs of each subject (which is 30).

2. Before decomposing the twelve multivariate signals by the NA-MEMD algorithm, each matrix was set up in a two-dimensional time series of the dimensions $[N_s \times N_e] \times N_t$. Therefore, the alignment among all IMFs across infants and over epochs was ascertained. A similar approach has been used previously by Hu and Liang (2011). **Figure 3** illustrates the decomposition process.

3. The resulting IMFs components after the decomposition process were slightly varied between channels. Since the EEG channel that yields the lowest number of IMFs upon decomposition gives ten modes, the first ten IMFs of each channel were considered for feature extraction. **Figure 4** depicts the sample of extracted IMFs from a channel that gave ten IMFs.

4. The frequencies of each IMF were then acquired by fast Fourier transform (FFT), and it was found that IMF1 to IMF3 are noisy and contain different oscillatory components. Thus, these modes were excluded from further analysis. IMF10 was also ignored as it represented the residue mode of some EEG channels, which might give unreal information about the signal. The scales of the remaining IMFs were localized approximately around the following ranges: IMF4 (15–26 Hz), IMF5 (10–13 Hz), IMF6 (6–8 Hz), IMF7 (3–4 Hz), IMF8 (1.5–3 Hz), and

IMF9 (0.5–1.5 Hz). Compared to five standard brain waves (Sanei and Chambers, 2007): IMF4 to IMF6 belong to beta, alpha, and theta bands, respectively; IMF7 to IMF9 all correspond to the delta band.

5. After selecting the IMFs, a dataset of the following dimensionality for each subject was achieved: $N_c \times N_i \times N_e \times N_t$, where N_c is the number of channels which is 12, N_i is the selected number of IMFs which is 6 (IMF4–IMF9), N_e is the number of epochs which is 30, and N_t is the number of the samples which is 1024.

Step 2: Feature Extraction
Weighted Phase-Lag Index

Before extracting the five graph-theoretical features, each IMF frequency scale alignment among channels was checked. Following that, the WPLI connectivity matrix was calculated between twelve channels for each IMF and each epoch. The generating WPLI matrices were averaged over the epochs, yielding one averaged connectivity matrix for each IMF. Then, each connectivity matrix was transformed into a complex connectivity network, and the graph-theoretical parameters were estimated to quantify its properties. The extracted graph parameters were then used to train and test the models.

Entropy Features

For each channel and each IMF, the proposed entropy measures were computed for each epoch. The calculated features were then averaged among the epochs to get one SampEn, one PEn,

FIGURE 3 | Proposed simultaneous decomposition method of the EEG signals. The data points of all infants (including all epochs) are stacked on top of each other. This process is done for each channel separately, ending up with 12 multivariate signals; each of them has the dimensionality of $[N_s \times N_e] \times N_t$, where N_s denotes the number of infants (which is 20), N_e is the number of epochs of each infant (which is 30), and N_t indicates the number of temporal samples (which is 1024). The NA-MEMD is then applied for each of the 12 multivariate signals separately.

FIGURE 4 | An example of a set of IMFs resulted from the NA-MEMD decomposition of the 2 s EEG signal. IMF1 to IMF3 considered noisy, and IMF10 represented the residue mode. IMF4 to IMF6 were localized in the beta, alpha, and theta bands, respectively, while IMF7 to IMF9 belonged to the delta band.

and one SpEn for each IMF signal. Thus, for each subject and each IMF signal, we ended up with 36 features (3 features × 12 channels). These features were then used to train and test the regression models.

Statistical Analysis: Correlation Coefficient

Correlation analysis was used to statistically measure the strength of the relationship between two random variables. There are several types of correlation methods, such as Pearson, Spearman, and Kendall. Pearson's correlation coefficient was adopted herein to evaluate whether there is a linear relationship between the proposed qEEG features and the cognitive scores. Although it is sensitive to outliers, Pearson's coefficient is chosen as it is the most widely used technique to measure the linear relationship between two variables, easy to compute and simple to interpret.

Pearson correlation coefficient, denoted by r, was adopted herein for this purpose. Theoretically, the value of r falls in the interval between +1 and −1, with 0 indicates no linear relationship, +1 refers to a perfect positive correlation, i.e., when one variable increases, the other increases too, while −1 indicates a perfect negative correlation, i.e., when one variable increases, the other decreases.

The significant test was conducted through the hypothesis test to evaluate the significance of the correlation coefficient. P-value assesses the null hypothesis that stated that there is no relationship between qEEG features and cognitive scores. The null hypothesis is successfully rejected when the p-value is below the significant level, which usually equals 0.05, indicating that the correlation coefficient result is statistically significant.

P-value is generally calculated based on t-statistic using the following equation:

$$ t = \frac{r \times \sqrt{n-2}}{\sqrt{1-r^2}}, \tag{10} $$

where r is the correlation coefficient, and n is the size of the dataset. Then the p-value is calculated from the t-distribution.

The correlation analysis was performed using MATLAB's statistics toolbox.

Thus, for the graph-theoretical of WPLI, the correlation coefficient was utilized to determine the correlation strength between the five graph-theoretical features and the cognitive scores in each IMF. The p-value was also calculated to identify the high influence predictors (qEEG features).

For entropy analysis, the correlation coefficient was performed to determine the linear dependency between each entropy feature

(SampEn, PEn, and SpEn) computed from each channel and the cognitive scores. The p-value was also computed to evaluate the statistically significant predictors. This process was repeated for each IMF separately.

Regression Model

A regression model was used to predict the later cognitive scores of the infants. The model tries to fit the relationship between the two proposed sets of EEG features (graph metrics and entropies) and the cognitive scores with the least possible error. The tree ensemble regression models were adopted in this study in order to reduce bias and variance in the imbalanced distribution of our dataset–as the distribution of the cognitive scores ranged between 74 and 145, such that most of the scores clustered above 95.

The basic idea of tree ensemble regression is using several combined models to obtain improved predictive performance (Moniz et al., 2017). Boosted trees regression and bagged trees regression were the two ensembles' models adopted in this study. Bagged tree regression randomly sampled the original dataset into different subsets with replacement. Several homogeneous models run independently on each subset in parallel, and the final predictive performance is obtained by combining the estimations of several models. In contrast, the boosted tree is a sequential ensemble method in which several homogenous models train adaptively. Each example in the dataset is assigned with weight. The examples that are incorrectly classified carry more weight than the examples that are correctly classified. Thus, the successor classifier focuses more on the example with the high weight that the predecessor classifier failed to classify correctly. A more detailed description of these models is available in Moniz et al. (2017) and MathWorks (2021).

Tree ensembles regression models were trained with both proposed sets of features separately. Regression learner apps within the statistics and machine learning toolbox in MATLAB was used to train the selected models.

Leave-one-subject-out cross-validation (LOSOCV) was used to assess the model performance in order to avoid the biased estimation of the prediction performance. This method works by splitting the dataset into two parts: one used for training and the other for testing. The training set contains N-1 individuals (where $N = 20$), and the remaining individual is turned into the testing set. Each individual is left out once in an iterative framework (N iterations), and then model statistics are evaluated by averaging the N independent regression outcomes.

The performance of regression models was evaluated by the traditional measures known as root mean square error (RMSE), mean absolute error (MAE), and R-squared. RMSE is the most frequently used metric. It refers to the square root of the average squared difference between the predicted score resulting from the regression model and the actual one. Lower RMSE indicates the better model's performance. MAE is the absolute difference between the predicted value and the target one, and as in the case of RMSE, the lowest value refers to the best model's performance. R-squared is another metric used to evaluate the performance of the regression model. It determines how well the model predicts the specific score by comparing the learned model with the constant baseline model. The constant baseline model is built by taking the mean of training data and drawing the line on the mean. The value of R-squared is usually less than or equal to 1 where the higher value refers to a better fit between predicted and actual values.

RESULTS

Weighted Phase-Lag Index-Based Functional Brain Connectivity Results

Table 3 shows the p-value results of the correlation analyses between graph-theoretical features and cognitive scores in each IMF component. In **Table 3**, features with the smallest p-value are shown with boldface, indicating the statistically significant correlation with the cognitive scores. These features were radius calculated from IMF7 (3–4 Hz) and transitivity, global efficiency, and characteristic path length computed from IMF8 (1.5–3 Hz). Correlation plots in **Figure 5** reveal that the radius and characteristic path length exhibit a significant negative correlation ($r = -0.46$, $p = 0.04$) and ($r = -0.45$, $p = 0.04$), respectively. Transitivity and global efficiency show a high positive correlation ($r = 0.48$, $p = 0.03$) and ($r = 0.49$, $p = 0.02$), respectively.

Considering that these features have highly significant correlation coefficients, they could greatly influence predicting the cognitive outcome. Thus, these four features were selected to be used in training and testing the regression models.

LOOCV was used to evaluate the models' performance to prevent potential bias from occurring due to overfitting. **Table 4** depicts the performance of tree ensemble regression models of the four selected features and their combinations.

TABLE 3 | P-values of the correlation analysis of the graph-theoretical features.

	IMF4 (15–26 Hz)	IMF5 (10–13 Hz)	IMF6 (6–8 Hz)	IMF7 (3–4 Hz)	IMF8 (1.5–3 Hz)	IMF9 (0.5–1.5 Hz)
Transitivity	0.12	0.93	0.33	0.62	**0.03**	0.99
Global efficiency	0.11	0.96	0.28	0.66	**0.02**	0.99
Radius	0.65	0.67	0.89	**0.04**	0.16	0.21
Diameter	0.26	0.63	0.4	0.76	0.76	0.9
Characteristic path length	0.18	0.87	0.43	0.54	**0.04**	0.9

Significant features are shown in boldface.
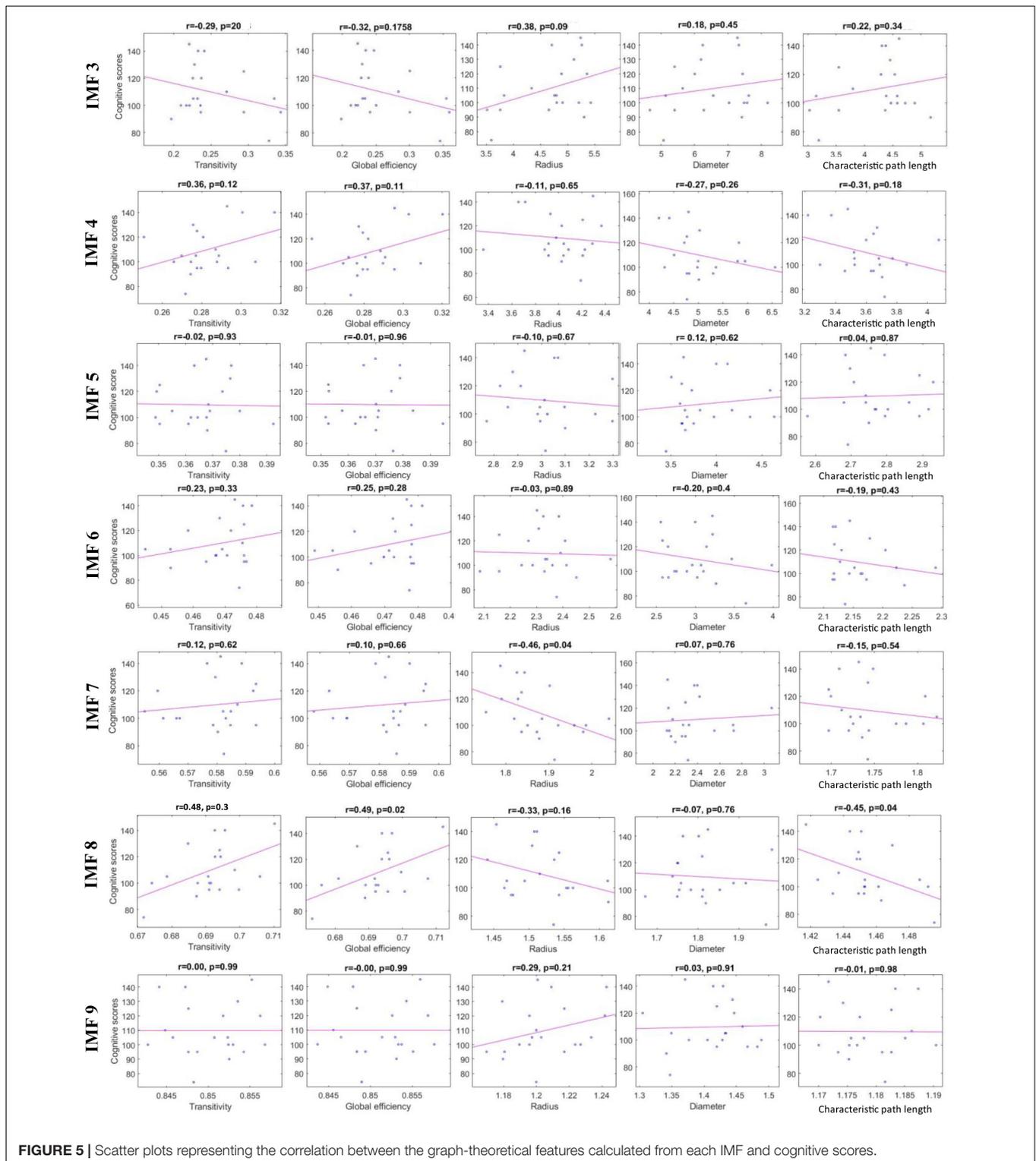
FIGURE 5 | Scatter plots representing the correlation between the graph-theoretical features calculated from each IMF and cognitive scores.

It is apparent from **Table 4** that the best performance–in terms of lowest RMSE, MSA, MAE, and highest R-squared–was achieved using radius network property from IMF7 (3–4 Hz). The visualization corresponding to this result represented by the difference between predicted scores and the actual scores is shown in **Figure 6**. The error rate between the predicted values and actual ones of majorities of the individual was acceptable as depicted in **Figure 6**. Other features such as transitivity, global efficiency and characteristic path length calculated from IMF8 (1.5–3 Hz) also give comparable results. This result implies that

TABLE 4 | Performance of the tree ensembles regression models using the significant graph-theoretical features.

Scale	Feature	RMSE	MAE	*R*-squared	Regression algorithm
IMF7 (3–4 Hz)	Radius	**16.775**	**12.7**	**0.24**	**Bagged trees**
		18.945	14.2	0.03	Boosted trees
IMF8 (1.5–3 Hz)	Transitivity	17.317	13.64	0.19	Bagged trees
		17.802	13.86	0.15	Boosted trees
	Global efficiency	17.26	13.64	0.2	Bagged trees
		17.71	13.82	0.15	Boosted trees
	Characteristic path length	16.98	13.28	0.22	Bagged trees
		17.78	13.85	0.15	Boosted trees
	Combination of transitivity, global efficiency, and characteristic path length	17.11	13.23	0.21	Bagged trees
		17.842	13.897	0.14	Boosted trees

The best model performance is shown in boldface.
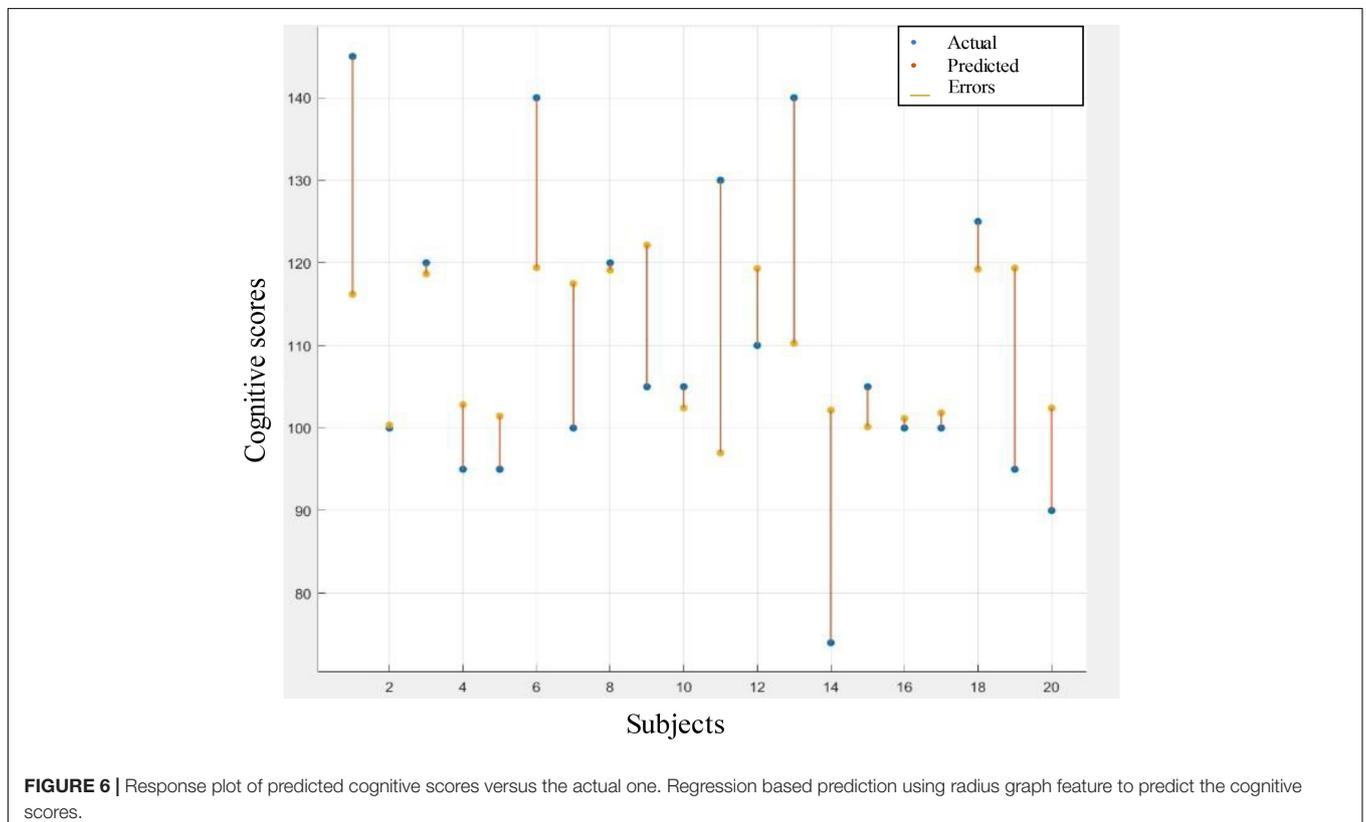
FIGURE 6 | Response plot of predicted cognitive scores versus the actual one. Regression based prediction using radius graph feature to predict the cognitive scores.

the network attributes–mainly radius–could provide valuable information regarding cognitive outcomes.

Entropy Analysis Results

Correlation analyses were carried out between the entropy measures, computed from each IMF and each channel, and the cognitive scores of all participants. Different embedding dimensions m, tolerances r, and time delay L were explored for SampEn and PEn estimations as suggested by Bruhn et al. (2000); Tudor et al. (2005); Olofsen et al. (2008). The correlation results indicate the robustness of SampEn and PEn with the small changes of the embedding parameters. **Table 5** presents the *p*-values of the correlation analyses using $m = 3$, $r = 1$, and $L = 1$ for SampEn computation, and $m = 3$ and $L = 1$ for PEn

calculation. The detailed correlation values of other embedding parameters are presented in the **Supplementary Table**.

The results presented in **Table 5** show that the significant *p*-values generally realize in the IMF9 (0.5–1.5 Hz) component of the left cerebral hemisphere. Particularly, the PEn calculated from channel C3 and SpEn computed from channels T3 and T5 exhibit significant correlations with the vector of the cognitive scores.

According to the correlation plots in **Figure 7**, the PEn feature of channel C3 and SpEn feature of channel T3 are shown to exhibit a significant negative correlation of ($r = -0.53$, $p = 0.01$) and ($r = -0.43$, $p = 0.05$), respectively. In addition, a high positive correlation is demonstrated by the SpEn of channel T5 ($r = 0.48$, $p = 0.03$). Thus, these features were selected to be used in the regression stage.

TABLE 5 | *P*-values of the correlation analysis of the complexity features.

	Channel index	IMF4 (15–26 Hz)	IMF5 (10–13 Hz)	IMF6 (6–8 Hz)	IMF7 (3–4 Hz)	IMF8 (1.5–3 Hz)	IMF9 (0.5–1.5 Hz)
SampEn	C3	0.79	0.97	0.95	0.56	0.51	0.19
	F3	0.42	0.86	0.77	0.28	0.40	0.66
	F7	0.86	0.41	0.95	0.99	0.83	0.50
	Fz	0.38	0.61	0.07	0.13	0.71	0.27
	O1	0.61	0.45	0.41	0.93	0.32	0.43
	O2	0.27	0.52	0.97	0.50	0.07	0.81
	P3	0.33	0.75	0.33	0.68	0.42	0.71
	P4	0.33	0.92	0.68	0.46	0.86	0.18
	T3	0.59	0.42	0.82	0.18	0.98	0.84
	T4	0.73	0.77	0.66	0.51	0.79	0.48
	T5	0.28	0.65	0.60	0.74	0.68	0.27
	T6	0.50	0.71	0.48	0.32	0.95	0.07
PEn	**C3**	0.68	0.80	0.30	0.79	0.27	**0.01**
	F3	0.32	0.52	0.69	0.24	0.54	0.27
	F7	0.27	0.48	0.42	0.56	0.63	0.12
	Fz	0.39	0.47	0.37	0.32	0.81	0.89
	O1	0.49	0.39	0.09	0.98	0.85	0.95
	O2	0.18	0.98	0.66	0.80	0.62	0.27
	P3	0.31	0.84	0.86	0.26	0.76	0.73
	P4	0.12	0.42	0.93	0.47	0.65	0.37
	T3	0.49	0.33	0.51	0.15	0.40	0.28
	T4	0.88	0.33	0.52	0.24	0.99	0.28
	T5	0.13	0.92	0.44	0.56	0.61	0.24
	T6	0.59	0.89	0.48	0.73	0.43	0.40
SpEn	C3	0.49	0.49	0.99	0.42	0.44	0.74
	F3	0.55	0.55	0.90	0.08	0.12	0.93
	F7	0.83	0.83	0.10	0.99	0.21	0.55
	Fz	0.56	0.56	0.30	0.42	0.85	0.69
	O1	0.56	0.56	0.39	0.90	0.56	0.57
	O2	0.96	0.96	0.94	0.16	0.14	0.85
	P3	0.40	0.40	0.69	0.53	0.89	0.12
	P4	0.72	0.72	0.57	0.69	0.55	0.47
	T3	0.60	0.60	0.74	0.26	0.62	**0.05**
	T4	0.65	0.65	0.78	0.88	0.24	0.73
	T5	0.69	0.69	0.75	0.14	0.26	**0.03**
	T6	0.78	0.78	0.68	0.96	0.73	0.11

Significant features are shown in boldface.

Table 6 illustrates the performance of LOSOCV when the models were trained and tested on the combination of the selected entropies. It can be noticed from the table that the boosted tree regression model achieved the best performance. **Figure 8** gives the visualization corresponding to this result. The figure reveals acceptable error rates, for most individuals, between the predicted values and the actual ones. This result indicates a reasonable prediction value for the cognitive scores using PEn and SpEn extracted from the IMF9 (0.5–1.5 Hz) component corresponding to the delta band. Further, we can infer that a good predictive value of the cognitive outcome is observed from the deterioration of EEG complexity produced by the left hemisphere.

DISCUSSION

The objective of the present investigation is to explore the effectiveness of employing qEEG analysis in the early prediction of cognitive outcome, assessed at 2 years of age following neonatal HIE. The early phase of a child's life is considered a critical stage for cognition, motor, language and social-emotional development owing to brain development and maturation of cortical architecture that are most rapidly established in this period (Ouyang et al., 2020). Early identification of the infants who have the cognitive impairment could help to provide a tailored intervention seeking to improve the outcome by utilization of this property of the brain which is called brain plasticity.

Two sets of features have been adopted for early identification of high-risk infants to develop cognitive impairment at 2 years of age, which were graph attributes of WPLI and complexity features extracted from EEG signals of twenty infants with neonatal HIE during their first week after birth.

The most significant challenge encountered in this study was that the distribution of the dataset was biased, with most cognitive scores clustered above 95, within 1 SD from the population mean. Most of the efforts in the machine learning community have been devoted to eliminating current challenges by designing an algorithm that can deal with bias and variance in the dataset. Tree ensembles regression, employed herein, is one of the efficient algorithms developed to handle this problem. It is designed to train multiple models and then combine their results to improve the performance of the final model. Furthermore, the subjective nature of the employed pre-processing techniques may also be considered as some artifacts still need to be removed using visual inspection. Hence, a fully automated process is of great interest to avoid subjective biases.

To the best of our knowledge, this research constitutes the first analysis on the impact of both proposed qEEG features calculated in the NA-MEMD domain for predicting cognitive outcome. Graph-theoretical metrics of WPLI were the first class of features adopted in this study. Statistical analysis shows a a significant correlation between the graph-theoretical features and the cognitive scores in the delta band connectivity corresponding to IMF7 (3–4 Hz) and IMF8 (1.5–3 Hz) components using radius, characteristic path length, transitivity, and global efficiency attributes. A strong negative relationship between radius and cognitive profiles is observed, indicating that the trend of the higher radius is correlated with poorer cognitive outcomes. This result suggests that the weak connections in brain networks (represented by increase in the radius) is negatively related to cognitive performance, i.e., increase in radius of the brain network is associated with poorer cognitive development.

On the other hand, a negative correlation is revealed in characteristic path length (a measure of network efficiency) with cognitive scores, displaying that an increase in characteristic path length is inversely associated with cognitive scores. This result indicates that a less efficient brain network in terms of global information transfer (measured by increase in characteristic path length) is negatively correlated with the cognitive level, i.e.,
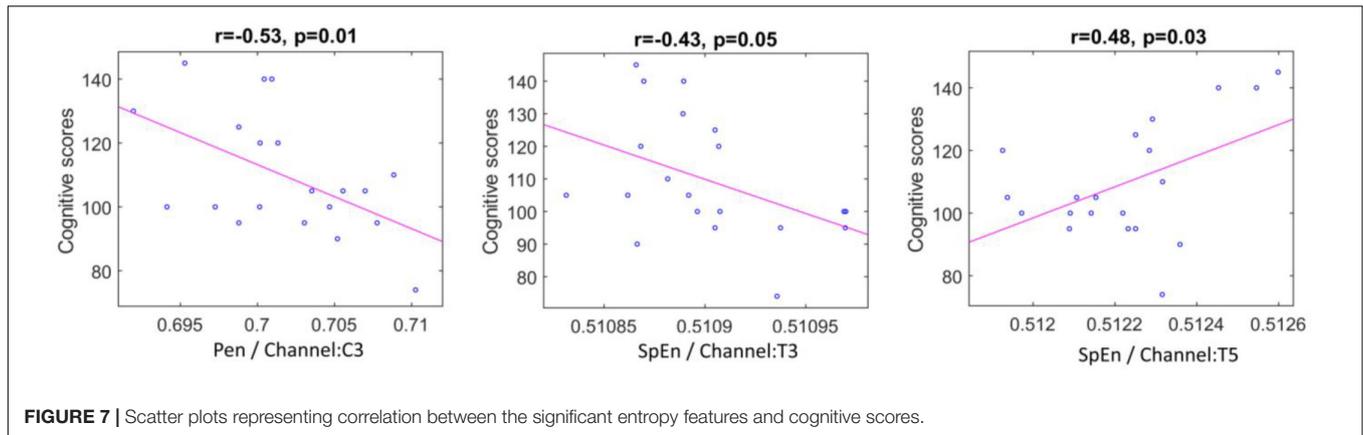
FIGURE 7 | Scatter plots representing correlation between the significant entropy features and cognitive scores.

increase in characteristic path length conducive to a reduction in the cognitive performance.

Moreover, the correlation analysis demonstrates a positive relationship between transitivity and global efficiency, and cognitive outcome, i.e., high transitivity and global efficiency, yielding higher cognitive scores. These results indicate that network efficiency in terms of information transfer between different brain regions is positively associated with cognitive outcome, i.e., increase in transitivity and global efficiency leads to better cognitive outcomes.

To sum up, the global graph-theoretical metrics (except diameter) appear promising to act as biomarkers for early prediction of cognitive outcome. This finding is consistent with existing studies other conditions. For example, the study by Peters et al. (2013) showed study showed increased characteristic path length and decreased global efficiency in brain networks of children with Autistic Spectrum Disorder (ASD).

The second class of features employed in this study were entropy measures, which are traditionally used to estimate the degree of EEG complexity. The existing entropy measures quantify the regularity of a time series represented on a single scale (Takahashi et al., 2010), Multiscale entropy (Costa et al., 2002) which measures the complexity considering different scales inherent in the signal. Though powerful, the multiscale entropy method is not well suited for studying non-linear

TABLE 6 | Performance of the tree ensemble regression models using the significant entropies features computed from IMF9 (0.5–1.5 Hz).

Features	RMSE	MAE	*R*-squared	Regression algorithm
PEn (C3)	17.069	14.025	0.21	Bagged tree
	16.856	14.151	0.23	Boosted tree
SpEn (T3)	17.993	15.116	0.13	Bagged tree
	17.876	13.823	0.14	Boosted tree
SpEn (T5)	18.789	14.26	0.05	Bagged tree
	19.957	15.058	0.07	Boosted tree
Combination of PEn (C3), SpEn (T3) and (T5)	16.283	12.655	0.29	Bagged tree
	14.271	**12.067**	**0.45**	**Boosted tree**

The best model performance is shown in boldface.

and non-stationary signals due to its linear extraction of scale (Peng et al., 2009). Subband wavelet entropy (SWE) (Al-Nashash et al., 2003) was also proposed to measure Shannon entropy from multiscale components. However, SWE is based on the wavelet method, which relies on predefined frequency ranges for the decomposition process. In the present work, we addressed these issues by computing the proposed entropies over different signal scales/IMFs extracted by NA-MEMD, which decomposes the signals adaptively and considering the non-linear and non-stationary nature of EEGs (Hu and Liang, 2011; Looney et al., 2015). The correlation analysis revealed a significant association in the delta band component, corresponding to IMF9, between entropies and the cognitive scores. Particularly, a significant positive correlation is observed between the SpEn extracted from the posterior brain area and the cognitive scores. This correlation indicates that the lower SpEn in this region was associated with a higher potential of cognitive impairments and vice versa. This result follows the general assumption of the physiological complexity reduction being related to various pathological processes (Catarino et al., 2011; Chu et al., 2017). A positive positive correlation between complexity measures and several cognitive functions at multiple brain regions was reported in different studies in older adults (Hu et al., 2016).

The analysis also demonstrates a negative correlation between the Pen, SpEn computed from the central brain region and the cognitive outcome. This result shows that the randomness behavior of the brain is negatively related to the level of cognitive function, i.e., an increase in the complexity of EEG signals leads to a reduction in the cognitive performance. This finding contradicts the general assumption of decreased complexity in a damaged brain. Li et al. (2010) reported a similar increase in the ApEn neurologically abnormal neonates with HIE or epilepsy compared to normal term neonates. Sajedi et al. (2013) also showed higher complexity in individuals with CP and reported that the CP neuronal networks functions are more random than in a normal brain. They indicated that this increase in complexity might result from reduced coordination among the neuronal regions in the disordered brain. This interpretation could be traced back to the non-linear dynamical theory, where it is known that non-linear coupled oscillations exhibit enhancement of complexity by the emergence of a chaotic state while reducing

FIGURE 8 | Response plot of predicted cognitive scores versus the actual one. Regression based prediction used the combination of the selected entropies features to predict the cognitive scores.

the coupled strength for the state of complete synchronization (Pikovsky et al., 2001). Thus, during decreasing coupled strength, the synchronization reduces, and the complexity enhances. Following this hypothesis, the decrease of entropy measures at the central region could be explained by the changes in the interactions of that brain area and other areas.

Nevertheless, the clinical significance of such inconsistent results remains unclear. Still, the findings indicate that the complexity measures of EEG could, at least, be useful in identifying abnormal brain function.

The correlational analyses also suggest that cognitive development of infants at the age of 2 years is associated with EEG complexity of the left hemisphere. Although several studies claimed that the right hemisphere is dominant in infants (Chiron, 1997; Adibpour et al., 2018), Chiron (1997) reported that the right hemisphere sustains the visuospatial abilities while the left hemisphere dominance for language function. Moreover, socio-cognitive function has also been linked by different studies to the neurophysiological processes of the left-brain. For example, there is ample evidence that relatively higher left than right frontal activity is related to social behavior (Harmon-Jones et al., 2010). Paulus et al. (2013) showed that prosocial understanding is associated with relatively stronger left frontal cortical activation in 2 years old infants. The study by Kühn-Popp et al. (2016) supported the socio-cognitive function of left-hemispheric brain maturation processes, which proved to be the prominent independent predictor of social communication abilities at 48 months.

Both proposed methods (WPLI and complexity analysis) reveal correlations between neonatal EEG characteristics and cognitive outcome in the delta band. This finding is consistent with Suppiej et al. (2017), where the authors concluded that the high value of the delta power spectral correlated with poor outcomes in preterm infants during the first year of age. Increased delta activity in EEG of children suffering from learning disorders was also confirmed by Martínez-Briones et al. (2020) and further support by Barttfeld et al. (2011) who reported the difference in delta band coherence between children having ASD and a control group. However, some studies have found the alterations in theta band waves in infants is a predictive biomarker for later cognitive impairment during the presenting of a specific stimulus. These studies suggested change in theta waves in infants when the subject tried to respond to an external stimulus. Jones et al. (2020) reported that EEG theta change in infancy during the presentation of dynamic movies of people and objects is a predictive biomarker for later cognitive deficits, particularly in high-risk populations. Braithwaite et al. (2020) also found increases in the theta wave recorded at 6-month-old infants at low-risk while presenting the non-social video. These changes in theta significantly predicted non-verbal cognitive ability measured at age 9-months. Begus et al. (2015) found that increases in frontal theta oscillations during object exploration correlated with subsequent recognition of that object in infants aged 11 months. The heterogeneity of the brain waves that play essential roles in the cognitive development of the infants can be attributed to the differences in the experimental

conditions. While most studies that found correlations between brain characteristics and cognitive functions in the delta band used resting-state EEG, the research that reported associations between EEG features and their corresponding cognitive profile in the theta wave have employed task-related EEG analysis.

Two tree ensembles regression models were explored to handle the bias distribution of our dataset. The significant graph-theoretical features (transitivity, global efficiency, radius and characteristic path length) and complexity features (SpEn and PEn) calculated from IMF components corresponding to the delta band were used to train and test the regression models. The best performance has been observed using boosted tree regression with a combination of Pen and SpEn features from left-brain regions (root mean square error score (RMSE) = 14.27, MAE = 12.07, and R-squared = 0.45). A comparable result was also observed using bagged tree regression with radius as a feature (RMSE = 16.78, MAE = 12.07, and R-squared = 0.24).

A key strength of this research was recognized when compared with the state-of-the-art of qEEG studies, shown in **Table 7**. To the best of our knowledge, our study is the first prospective study to date performed in neonates (in the first week of birth) investigating early non-linear qEEG characteristics (WPLI and complexity measures) and their prognostic value for cognitive outcome at 24 months of age. Furthermore, all studies existing in literature have used the linear qEEG such as coherence (Kühn-Popp et al., 2016), EEG continuity (West et al., 2005), and spectral power (Suppiej et al., 2017; Cainelli et al., 2021), which may not be optimal to capture the complex characteristics of the EEG spectra. Non-linear methods adopted in this study provided deep insight into the underlying brain functions and dynamics.

TABLE 7 | Comparison of the qEEG state-of-the-art methods employed for predicting cognitive outcomes.

Author	Dataset	Features	Evaluation methods	Outcome assessment	Findings
Lloyd et al., 2021	57 preterm infants	EEG grading	Spearman's correlation coefficient	BSITD-III	Moderate to large negative correlation between EEG grade and Bayley-III subscale
Suppiej et al., 2017	21 preterm infants	Power spectral analysis	Spearman's correlation coefficient	Griffiths Scale of Mental Development	Negative correlation between the delta spectral power and Griffiths scores developmental quotients ($r = -0.68$, $p = 0.015$). Positive correlation between alpha and beta power spectral and Griffiths developmental quotients ($r = 0.61$, $p = 0.032$).
Cainelli et al., 2021	26 preterm infants	Power spectral analysis	Bayesian correlation	Wechsler Pre-school and Primary Scale of Intelligence III (WPPSI-III) test	Significant association between spectral frequency bands and visual and auditory attention tests.
West et al., 2005	44 preterm infants	EEG continuity	Linear regression	BSITD-II	Significant correlation between mental developmental indices and continuity feature of EEG at different amplitude setting: 10 and 25 μV thresholds (R-squared = 0.19, $P = 0.0032$ and R-squared = 0.10, $p = 0.04$, respectively).
Kühn-Popp et al., 2016	32 infants	EEG coherence measures	Linear regression	Coding-scheme for mental state terms	Significant correlation between left hemisphere coherence and epistemic language at 48 months ($r = 0.59$, $p = 0.003$). Regression analyses showed, left-coherence scores are the most important predictor of epistemic state talk at 48 months
Current study	20 infants born with HIE	• Graph-theoretical features derived from WPLI • Entropy features	• Pearson linear correlation coefficient • Set of regression models	BSITD-III	**Connectivity:** significant correlation between transitivity, global efficiency, radius, and characteristic path length and cognitive outcomes. Reasonable regression performance with radius feature: RMSE (16.775), MAE (12.702), and R-square (0.24). **Complexity:** significant correlation between PEn and SpEn measured from left regions of brain and cognitive profile. Good regression performance with combination of PEn from C3 and SpEn T3 and T5: RMSE (14.27), MAE (12.07), and R-square (0.45).

Several studies investigated the frequency activity underlying EEG by spectrally analyzing the signal using methods that rely on the predefined frequency of traditional brainwaves, as in Kühn-Popp et al. (2016), Suppiej et al. (2017), and Cainelli et al. (2021). Prior selection of frequency ranges may result in potentially informative brain waves being missed, specifically in the case of infants, due to the well-known variability between them and older individuals in the neural oscillations of interest. Moreover, the predefined basis may not be able to fit all the non-linear and non-stationary phenomena (Huang et al., 1998). This constraint has been settled in our proposed approach using the NA-MEMD method, which decomposes the signals adaptively. Thus, with this method, we ascertained that all meaningful brain dynamics are included in the analysis, and no misleading energy-frequency distribution will result from analyzing the non-stationary and non-linear signals. Nevertheless, the advantages of the EMD-based methods have a price of being empirically, not theoretically, defined. In addition, the NA-MEMD method has a very high computational complexity having a subspace of multivariate independent white noise equal to the original multivariate signal.

In summary, this study provided a regression-based machine learning framework to objectively predict cognitive outcome at toddler age, and a good performance was achieved. We found that the global network attributes and entropy features could serve as early markers for cognitive development. Further, the experimental results found a association between neonate's left-brain region and cognitive deficit emerging at 2 years. Due to the limited dataset size, this study can only be considered a proof of concept work. It is provided the initial proof of that employing the qEEG features within regression-based machine learning frameworks could capture the individual variability inherited in infants' developing brains. This study lays the groundwork for future investigations into using these features as the potential biomarkers for predicting cognitive development in a number of populations who are at risk for long term cognitive impairment or intellectual disability. This could assist in establishing tailored intervention programmes at an early stage to improve outcomes. Nevertheless, a further refinement of the proposed analysis with larger sample sizes is required to validate the findings.

CONCLUSION

In this investigation, the analysis of qEEG successfully predicted the cognitive outcome at 2 years of age in infants with neonatal HIE. Complexity analysis (SampEn, PEn, and SpEn) and functional brain connectivity (WPLI) measures were evaluated by correlation analysis and regression-based machine learning framework. Pearson linear correlation analysis showed a strong correlation between graph-theoretical features of WPLI, PEn,

SpEn, and cognitive scores. The tree ensembles regression models have achieved comparable performance in both methods, with relative superiority of complexity measures using combination between PEn and SpEn; RMSE (14.27), MAE (12.07), and R-square (0.45). Therefore, the findings of this study have provided insight into the possibility of using graph-theoretical features and entropy measures derived from the delta band as biomarkers for early prediction of cognitive development.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by HRA and Health and Care Research Wales, HCRW (Reference ID 20/HRA/0260; IRAS project ID 278072; University Hospital Southampton R&D protocol number RHM CHI1047). Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

NA: methodology, software, formal analysis, data curation, writing–original draft, and visualization. DB: methodology, software, formal analysis, data curation, and writing–original draft. BV: providing the neurological evaluation information. DK: providing the EEG dataset. KM: conceptualization, validation, investigation, resources, and supervision. All authors contributed to the article and approved the submitted version.

FUNDING

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnhum.2021.795006/full#supplementary-material

REFERENCES

Adibpour, P., Dubois, J., and Dehaene-Lambertz, G. (2018). Right but not left hemispheric discrimination of faces in infancy. *Nat. Hum. Behav.* 2, 67–79. doi: 10.1038/s41562-017-0249-4

Al-Nashash, H. A., Paul, J. S., Ziai, W. C., Hanley, D. F., and Thakor, N. V. (2003). Wavelet entropy for subband segmentation of EEG during injury and recovery. *Ann. Biomed. Eng.* 31, 653–658. doi: 10.1114/1.1575757

Apicella, F., Sicca, F., Federico, R. R., Campatelli, G., and Muratori, F. (2013). Fusiform gyrus responses to neutral and emotional faces in children with autism

spectrum disorders: a high density ERP study. *Behav. Brain Res.* 251, 155–162. doi: 10.1016/j.bbr.2012.10.040

Babiloni, C., Brancucci, A., Vecchio, F., Arendtnielsen, L., Chen, A., and Rossini, P. (2006). Anticipation of somatosensory and motor events increases centro-parietal functional coupling: an EEG coherence study. *Clin. Neurophysiol.* 117, 1000–1008. doi: 10.1016/j.clinph.2005.12.028

Bakheet, D., Alotaibi, N., Konn, D., Vollmer, B., and Maharatna, K. (2021). Prediction of cerebral palsy in newborns with hypoxic-ischemic encephalopathy using multivariate EEG analysis and machine learning. *IEEE Access* 9, 137833–137846.

Bandt, C., and Pompe, B. (2002). Permutation entropy: a natural complexity measure for time series. *Phys. Rev. Lett.* 88:174102. doi: 10.1103/PhysRevLett.88.174102

Barttfeld, P., Wicker, B., Cukier, S., Navarta, S., Lew, S., and Sigman, M. (2011). A big-world network in ASD: dynamical connectivity analysis reflects a deficit in long-range connections and an excess of short-range connections. *Neuropsychologia* 49, 254–263. doi: 10.1016/j.neuropsychologia.2010.11.024

Begus, K., Southgate, V., and Gliga, T. (2015). Neural mechanisms of infant learning: Differences in frontal theta activity during object exploration modulate subsequent object recognition. *Biol. Lett.* 11, 3–4. doi: 10.1098/rsbl.2015.0041

Braithwaite, E. K., Jones, E. J. H., Johnson, M. H., and Holmboe, K. (2020). Dynamic modulation of frontal theta power predicts cognitive ability in infancy. *Dev. Cogn. Neurosci.* 45:100818. doi: 10.1016/j.dcn.2020.100818

Breakspear, M., Williams, L. M., and Stam, C. J. (2004). A novel method for the topographic analysis of neural activity reveals formation and dissolution of 'dynamic cell assemblies'. *J. Comput. Neurosci.* 16, 49–68. doi: 10.1023/b:jcns.0000004841.66897.7d

Bruhn, J., Röpcke, H., and Hoeft, A. (2000). approximate entropy as an electroencephalographic measure of anesthetic drug effect during desflurane anesthesia. *Anesthesiology* 92, 715–726. doi: 10.1097/00000542-200003000-00016

Bruña, R., Maestú, F., and Pereda, E. (2018). Phase locking value revisited: teaching new tricks to an old dog. *J. Neural Eng.* 15:056011. doi: 10.1088/1741-2552/aacfe4

Byeon, J. H., Kim, G.-H., Kim, J. Y., Sun, W., Kim, H., and Eun, B.-L. (2015). Cognitive dysfunction and hippocampal damage induced by hypoxic-ischemic brain injury and prolonged febrile convulsions in immature rats. *J. Korean Neurosurg. Soc.* 58:22. doi: 10.3340/jkns.2015.58.1.22

Cainelli, E., Vedovelli, L., Wigley, I. L. C. M., Bisiacchi, P. S., and Suppiej, A. (2021). Neonatal spectral EEG is prognostic of cognitive abilities at school age in premature infants without overt brain damage. *Eur. J. Pediatr.* 180, 909–918. doi: 10.1007/s00431-020-03818-x

Catarino, A., Churches, O., Baron-Cohen, S., Andrade, A., and Ring, H. (2011). Atypical EEG complexity in autism spectrum conditions: a multiscale entropy analysis. *Clin. Neurophysiol.* 122, 2375–2383. doi: 10.1016/j.clinph.2011.05.004

Chaumon, M., Bishop, D. V. M., and Busch, N. A. (2015). A practical guide to the selection of independent components of the electroencephalogram for artifact correction. *J. Neurosci. Methods* 250, 47–63. doi: 10.1016/j.jneumeth.2015.02.025

Chiron, C. (1997). The right brain hemisphere is dominant in human infants. *Brain* 120, 1057–1065. doi: 10.1093/brain/120.6.1057

Chu, Y. J., Chang, C. F., Shieh, J. S., and Lee, W. T. (2017). The potential application of multiscale entropy analysis of electroencephalography in children with neurological and neuropsychiatric disorders. *Entropy* 19, 1–13. doi: 10.3390/e19080428

Costa, M., Goldberger, A. L., and Peng, C.-K. (2002). Multiscale entropy analysis of complex physiologic time series. *Phys. Rev. Lett.* 89:068102.

Cui, D., Wang, J., Bian, Z., Li, Q., Wang, L., and Li, X. (2015). Analysis of entropies based on empirical mode decomposition in amnesic mild cognitive impairment of diabetes mellitus. *J. Innov. Opt. Health Sci.* 08:1550010.

David, O., Cosmelli, D., and Friston, K. J. (2004). Evaluation of different measures of functional connectivity using a neural mass model. *Neuroimage* 21, 659–673. doi: 10.1016/j.neuroimage.2003.10.006

Gabor, D. (1946). Theory of communication. Part 1: the analysis of information. *J. Inst. Electr. Eng. Part III Radio Commun. Eng.* 93, 429–441.

Harmon-Jones, E., Gable, P. A., and Peterson, C. K. (2010). The role of asymmetric frontal cortical activity in emotion-related phenomena: a review and update. *Biol. Psychol.* 84, 451–462. doi: 10.1016/j.biopsycho.2009.08.010

He, L., Li, H., Holland, S. K., Yuan, W., Altaye, M., and Parikh, N. A. (2018). Early prediction of cognitive deficits in very preterm infants using functional connectome data in an artificial neural network framework. *Neuroimage Clin.* 18, 290–297. doi: 10.1016/j.nicl.2018.01.032

Hu, C.-J., Chung, C.-C., and Kang, J.-H. (2016). Measuring entropy in functional neuroscience: pathophysiological and clinical applications. *Neurosci. Neuroecon.* 5, 45–53.

Hu, M., and Liang, H. (2011). Intrinsic mode entropy based on multivariate empirical mode decomposition and its application to neural data analysis. *Cogn. Neurodyn.* 5, 277–284. doi: 10.1007/s11571-011-9159-8

Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., et al. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. R. Soc. London. Ser. A Math. Phys. Eng. Sci.* 454, 903–995. doi: 10.21105/joss.02977

Jones, E. J. H., Goodwin, A., Orekhova, E., Charman, T., Dawson, G., Webb, S. J., et al. (2020). Infant EEG theta modulation predicts childhood intelligence. *Sci. Rep.* 10:11232. doi: 10.1038/s41598-020-67687-y

Kong, A. H. T., Lai, M. M., Finnigan, S., Ware, R. S., Boyd, R. N., and Colditz, P. B. (2018). Background EEG features and prediction of cognitive outcomes in very preterm infants: a systematic review. *Early Hum. Dev.* 127, 74–84. doi: 10.1016/j.earlhumdev.2018.09.015

Kühn-Popp, N., Kristen, S., Paulus, M., Meinhardt, J., and Sodian, B. (2016). Left hemisphere EEG coherence in infancy predicts infant declarative pointing and preschool epistemic language. *Soc. Neurosci.* 11, 49–59. doi: 10.1080/17470919.2015.1024887

Kułak, W., and Sobaniec, W. (2003). Spectral analysis and EEG coherence in children with cerebral palsy: spastic diplegia. *Przegl. Lek.* 60, 23–27.

Li, L., Chen, W., Shao, X., and Wang, Z. (2010). Analysis of amplitude-integrated EEG in the newborn based on approximate entropy. *IEEE Trans. Biomed. Eng.* 57, 2459–2466. doi: 10.1109/TBME.2010.2055863

Li, X., Li, D., Liang, Z., Voss, L. J., and Sleigh, J. W. (2008). Analysis of depth of anesthesia with Hilbert–Huang spectral entropy. *Clin. Neurophysiol.* 119, 2465–2475. doi: 10.1016/j.clinph.2008.08.006

Lloyd, R. O., O'Toole, J. M., Livingstone, V., Filan, P. M., and Boylan, G. B. (2021). Can EEG accurately predict 2-year neurodevelopmental outcome for preterm infants? *Arch. Dis. Child. Fetal Neonatal Ed.* 106, 535–541. doi: 10.1136/archdischild-2020-319825

Looney, D., Hemakom, A., and Mandic, D. P. (2015). Intrinsic multi-scale analysis: a multi-variate empirical mode decomposition framework. *Proc. R. Soc. A Math. Phys. Eng. Sci.* 471:20140709. doi: 10.1098/rspa.2014.0709

Mallat, S. G. (1989). A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Trans. Pattern Anal. Mach. Intell.* 11, 674–693.

Martínez-Briones, B., Fernández-Harmony, T., Garófalo Gómez, N., Biscay-Lirio, R., and Bosch-Bayard, J. (2020). Working memory in children with learning disorders: an EEG power spectrum analysis. *Brain Sci.* 10, 817.

MathWorks (2021). *Regression Learner App, The MathWorks, Inc.* [Online]. Available online at: https://www.mathworks.com/help/stats/regression-learner-app.html?s_tid=CRUX_lftnav (accessed February 20, 2021)

Mika (2010). *Brain Connectivity Toolbox, Mathworks* [Online]. Available: https://sites.google.com/site/bctnet/ (accessed April 20, 2021).

Moeskops, P., Išgum, I., Keunen, K., Claessens, N. H. P., van Haastert, I. C., Groenendaal, F., et al. (2017). Prediction of cognitive and motor outcome of preterm infants based on automatic quantitative descriptors from neonatal MR brain images. *Sci. Rep.* 7:2163. doi: 10.1038/s41598-017-02307-w

Moniz, N., Branco, P., Torgo, L., and Krawczyk, B. (2017). Evaluation of ensemble methods in imbalanced regression tasks. *Proc. Mach. Learn. Res.* 74, 129–140. doi: 10.1016/j.jbi.2004.07.008

Olofsen, E., Sleigh, J. W., Dahan, A., and Zealand, N. (2008). Permutation entropy of the electroencephalogram: a measure of anaesthetic drug effect. *Br. J. Anaesth.* 101, 810–821. doi: 10.1093/bja/aen290

Ouyang, M., Peng, Q., Jeon, T., Heyne, R., Chalak, L., and Huang, H. (2020). Diffusion-MRI-based regional cortical microstructure at birth for predicting neurodevelopmental outcomes of 2-year-olds. *Elife* 9:e58116. doi: 10.7554/eLife.58116

Paulus, M., Kühn-Popp, N., Licata, M., Sodian, B., and Meinhardt, J. (2013). Neural correlates of prosocial behavior in infancy: different neurophysiological mechanisms support the emergence of helping and comforting. *Neuroimage* 66, 522–530. doi: 10.1016/j.neuroimage.2012.10.041

Peng, C., Costa, M., and Goldberger, A. (2009). Adaptive data analysis of complex fluctuations in physiologic time series. *Adv. Adapt. Data Anal.* 1, 61–70. doi: 10.1142/S1793536909000035

Peters, J. M., Taquet, M., Vega, C., Jeste, S. S., Fernández, I. S., Tan, J., et al. (2013). Brain functional networks in syndromic and non-syndromic autism: a graph theoretical study of EEG connectivity. *BMC Med.* 11:54. doi: 10.1186/1741-7015-11-54

Pikovsky, A., Rosenblum, M., and Kurths, J. (2001). *Synchronization A Universal Concept in Nonlinear Sciences*. New York, NY: Cambridge University Press.

Pincus, S. M. (1991). Approximate entropy as a measure of system complexity. *Proc. Natl. Acad. Sci. U.S.A.* 88, 2297–2301.

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

Saby, J. N., and Marshall, P. J. (2012). The utility of EEG band power analysis in the study of infancy and early childhood. *Dev. Neuropsychol.* 37, 253–273. doi: 10.1080/87565641.2011.614663

Sajedi, F., Ahmadlou, M., Vameghi, R., Gharib, M., and Hemmati, S. (2013). Linear and nonlinear analysis of brain dynamics in children with cerebral palsy. *Res. Dev. Disabil.* 34, 1388–1396. doi: 10.1016/j.ridd.2013.01.016

Sakkalis, V. (2011). Review of advanced techniques for the estimation of brain connectivity measured with EEG/MEG. *Comput. Biol. Med.* 41, 1110–1117. doi: 10.1016/j.compbiomed.2011.06.020

Sanei, S., and Chambers, J. (2007). *EEG Signal Processing*. New York, NY: John Wiley & Sons.

Slaughter, L. A., Bonfante-Mejia, E., Hintz, S. R., Dvorchik, I., and Parikh, N. A. (2016). Early conventional MRI for prediction of neurodevelopmental impairment in extremely-low-birth-weight infants. *Neonatology* 110, 47–54. doi: 10.1159/000444179

Stam, C. J., Nolte, G., and Daffertshofer, A. (2007). Phase lag index: assessment of functional connectivity from multi channel EEG and MEG with diminished bias from common sources. *Hum. Brain Mapp.* 28, 1178–1193. doi: 10.1002/hbm.20346

Stam, C. J., van der Made, Y., Pijnenburg, Y. A. L., and Scheltens, P. (2003). EEG synchronization in mild cognitive impairment and Alzheimer's disease. *Acta Neurol. Scand.* 108, 90–96.

Sui, J., Jiang, R., Bustillo, J., and Calhoun, V. (2020). Neuroimaging-based individualized prediction of cognition and behavior for mental disorders and health: methods and promises. *Biol. Psychiatry* 88, 818–828. doi: 10.1016/j.biopsych.2020.02.016

Suppiej, A., Cainelli, E., Cappellari, A., Trevisanuto, D., Balao, L., Di Bono, M. G., et al. (2017). Spectral analysis highlight developmental EEG changes in preterm infants without overt brain damage. *Neurosci. Lett.* 649, 112–115. doi: 10.1016/j.neulet.2017.04.021

Sweeney-Reed, C. M., and Nasuto, S. J. (2007). A novel approach to the detection of synchronisation in EEG based on empirical mode decomposition. *J. Comput. Neurosci.* 23, 79–111. doi: 10.1007/s10827-007-0020-3

Takahashi, T., Cho, R. Y., Mizuno, T., Kikuchi, M., Murata, T., Takahashi, K., et al. (2010). Antipsychotics reverse abnormal EEG complexity in drug-naive schizophrenia: a multiscale entropy analysis. *Neuroimage* 51, 173–182. doi: 10.1016/j.neuroimage.2010.02.009

Tudor, M., Tudor, L., and Tudor, K. I. (2005). [Hans berger (1873-1941)–the history of electroencephalography]. *Acta Med. Croatica* 59, 307–313.

ur Rehman, N., and Mandic, D. P. (2011). Filter bank property of multivariate empirical mode decomposition. *IEEE Trans. Signal Process.* 59, 2421–2426.

ur Rehman, N., Xia, Y., and Mandic, D. P. (2010). "Application of multivariate empirical mode decomposition for seizure detection in EEG signals," in *Proceedings of the 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*, Buenos Aires, 1650–1653. doi: 10.1109/IEMBS.2010.5626665

Vecchio, F., Miraglia, F., and Maria Rossini, P. (2017). Connectome: graph theory application in functional brain network architecture. *Clin. Neurophysiol. Pract.* 2, 206–213. doi: 10.1016/j.cnp.2017.09.003

Vinck, M., Oostenveld, R., van Wingerden, M., Battaglia, F., and Pennartz, C. M. A. (2011). An improved index of phase-synchronization for electrophysiological data in the presence of volume-conduction, noise and sample-size bias. *Neuroimage* 55, 1548–1565. doi: 10.1016/j.neuroimage.2011.01.055

West, C. R., Battin, M. R., Williams, C. E., Dezoete, J. A., and Harding, J. E. (2005). 413 early quantitative electroencephalographic measures of continuity are associated with neurodevelopmental outcome at 18 months in preterm infants. *Pediatr. Res.* 58, 425–425.

Zahra, A., Kanwal, N., Ehsan, S., and Mcdonald-maier, K. D. (2010). *Introduction to Pattern Recognition*. Amsterdam: Elsevier, 88.

Zhao, T., Xu, Y., and He, Y. (2019). Graph theoretical modeling of baby brain networks. *Neuroimage* 185, 711–727. doi: 10.1016/j.neuroimage.2018.06.038