Evaluating Complexity of Fetal MEG Signals A Comparison of Different Metrics and Their Applicability

Julia Moser, Siouar Bensaid, Eleni Kroupi, Franziska Schleger, Fabrice Wendling, Giulio Ruffini, Hubert Preißl

To cite this version:

Julia Moser, Siouar Bensaid, Eleni Kroupi, Franziska Schleger, Fabrice Wendling, et al.. Evaluating Complexity of Fetal MEG Signals A Comparison of Different Metrics and Their Applicability. Frontiers in Systems Neuroscience, Frontiers, 2019, 13, pp.23. 10.3389/fnsys.2019.00023. hal-02177170

HAL Id: hal-02177170
https://hal-univ-rennes1.archives-ouvertes.fr/hal-02177170

Submitted on 4 Sep 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Evaluating Complexity of Fetal MEG Signals: A Comparison of Different Metrics and Their Applicability

Julia Moser1*, Siouar Bensaid2, Eleni Kroupi3, Franziska Schleger1, Fabrice Wendling2, Giulio Ruffini3 and Hubert Preißl1

1fMEG Center/Internal Medicine IV/Institute for Diabetes Research and Metabolic Diseases of the Hemholtz Center Munich at the University of Tübingen, Tübingen, Germany, 2INSERM, U1099, Rennes 1 University, Rennes, France, 3Starlab SLU, Barcelona, Spain

In this work, we aim to investigate whether information based metrics of neural activity are a useful tool for the quantification of consciousness before and shortly after birth. Neural activity is measured using fetal magnetoencephalography (fMEG) in human fetuses and neonates. Based on recent theories on consciousness, information-based metrics are established to measure brain complexity and to assess different levels of consciousness. Different metrics (measures of entropy, compressibility and fractality) are, thus, explored in a reference population and their usability is evaluated. For comparative analysis, two fMEG channels were selected: one where brain activity was previously detected and one at least 15 cm away, that represented a control channel. The usability of each metric was evaluated and results from the brain and control channel were compared. Concerning the ease of use with fMEG data, Lempel-Ziv-Complexity (LZC) was evaluated as best, as it is unequivocal and needs low computational effort. The fractality measures have a high number of parameters that need to be adjusted prior to analysis and therefore forfeit comparability, while entropy measures require a higher computational effort and more parameters to adjust compared to LZC. Comparison of a channel with brain activity and a control channel in neonatal recordings showed significant differences in most complexity metrics. This clear difference can be seen as proof of concept for the usability of complexity metrics in fMEG. For fetal data, this comparison produced less clear results which can be related to leftover maternal signals included in the control channel. Further work is necessary to conclusively interpret results from the analysis of fetal recordings. Yet this study shows that complexity metrics can be used for fMEG data on early consciousness and the evaluation gives a guidance for future work. The inconsistency of results from different metrics highlights the challenges of working with complexity metrics as neural correlates of consciousness, as well as the caution one should apply to interpret them.

Keywords: fMEG, complexity, LZC, entropy, fractality, neurodevelopment
INTRODUCTION

Consciousness is known to be one of the characteristics that make humans unique. But when does this aspect of the human mind arise? Is it possible that consciousness already exists before birth? From the 24th week of gestation, a fetus can process sensory stimuli at a cortical level, as thalamocortical connections are already established (Kostović and Judaš, 2010). Long range pyramidal neurons—which are known to be important for conscious processing (Dehaene et al., 1998)—are developed around week 26 (Lagarrucranta and Changeux, 2009). Yet, it is difficult to assess the conscious state of a fetus in the mother’s womb. Fetal magnetoencephalography (fMEG) is a tool to noninvasively investigate fetal brain activity in the last trimester of pregnancy and in neonates shortly after birth (Preissl et al., 2004). This measurement of fetal/neonatal brain activity makes it possible to investigate neural correlates of consciousness and pursue the question of the debut of consciousness in human life.

During the last decades, work in the field of disorders of consciousness led to an increased interest in neural correlates of consciousness. Neural processes that allow conscious experience need to be highly integrated and differentiated, which are properties that can be measured by complexity of neurological data (Tononi and Edelman, 1998). In nature, complexity of physiology is related to the adaptive capacity of an organism (Costa et al., 2002). This is translated into physiological signals with long-range correlations across various spatio-temporal scales—a behavior that is named self-organization—that indicate the presence of self-invariant and self-similar structures (Pritchard and Duke, 1995). The self-organizational properties of a complex system can be quantified by estimating its dimension (Théiler, 1990), or its ability to compress information (Cover and Thomas, 2012; Ruffini, 2017a).

For a system to be complex, it has to operate on several scales and also show an interplay between those scales (Lutzenberger et al., 1995). This is a property of a so-called chaotic system and can be measured in space and time (Elbert et al., 1994). Typically chaos in space is estimated with the fractal dimension, which is defined as the dimension of a strange attractor towards which a complex system evolves in phase-space (Grassberger and Procaccia, 1983b). Thus, the fractal dimension, namely D, describes the overall complexity of an object, which can be the geometrical complexity, the space filling property, the roughness of a surface, or the variation of a time series. The fractal dimension D is defined by the logarithmic ratio of change in detail with change in scale (Di Ieva, 2016). This relates to the distinct characteristic of a fractal, namely the property of self-similarity: i.e., pieces of an object are similar to larger pieces of it as well as to the whole object (Eke et al., 2002). In nature, fractals are usually only statistically self-similar which means that smaller excerpts are not necessarily exact copies of the larger ones, but they are the same on average (Pritchard and Duke, 1995). In contrast to fractal dimension, entropy gives information about the dynamical properties of an attractor and not about its geometrical shape (Rodríguez-Bermúdez and García-Laencina, 2015). Chaos in time, therefore, relates to this stability and sensitivity to initial conditions (Elbert et al., 1994; Baranger, 2000). Related to those properties of complex systems, several measures were developed to quantify their complexity. The sensitivity to initial conditions can be quantified in terms of the Lyapunov exponent and the Kolmogorov entropy, also known as information dimension (Théiler, 1990). Different entropy measures, as well as the measure of compressibility, can be employed for this. Ruffini (2017a) recently proposed a theory of consciousness that considers the brain an engine that strives to model the world with simplicity, while learning it is a result of exchanging information with it. According to this theory, the ability of the brain to compress information is an indicator of consciousness.

In consciousness neuroscience, this quantification of complexity is used in different scenarios. The main areas of application are research with patients with disorders of consciousness, anesthesia monitoring and sleep studies (e.g., Casarotto et al., 2016). For instance, Burioka et al. (2005) showed that approximate entropy calculated on small segments of electroencephalography (EEG) data decreases with depth of sleep. In particular, there is a linear decrease from wakefulness over sleep stage 1 until 4, while rapid eye movement sleep showed values similar to wakefulness. Analysis of data from EEG, MEG and intracranial EEG recordings confirmed this drop of complexity with a drop in wakefulness (Mateos et al., 2018). Complexity was calculated with entropy and compressibility measures. Zhang et al. (2009) could differentiate between active sleep and quiet sleep in newborns by means of sample entropy (SE). Similarly, the dimensional complexity of the EEG pattern measured by correlation dimension (CD) was found to be higher in active sleep compared to quiet sleep for infants in their 1st months of life (Janjarasjitt et al., 2008b; Scher et al., 2009). Furthermore, scale-free properties caused by the self-similarity of fractals can be used to differentiate between sleep stages whereas an increase or decrease of values depends on the scale-free parameter estimated (Weiss et al., 2009).

In general, the measurement of this scale-free behavior appears promising in the investigation of state transitions (Weiss et al., 2009). Studies with Propofol anesthesia showed a change in scale-free behavior before and after loss of consciousness (Eagleman et al., 2018) and a difference between wakefulness and loss of consciousness as well as recovery from anesthesia (Tagliazucchi et al., 2016). Also with the help of entropy measures, Eagleman et al. (2018) could show a change in complexity of scalp EEG data before and after loss of responsiveness in anesthesia patients. Similarly, Schartner et al. (2015) could distinguish between loss of consciousness during Propofol-induced anesthesia and wakeful rest by means of entropy measures as well as compressibility measures. Furthermore, higher entropy values were shown in the EEG data of healthy control participants compared to unresponsive wakefulness patients, matched for sex and age (Sără and Pistoia, 2010).

Compressibility measures in combination with transcranial magnetic stimulation are widely used to distinguish between patients with different disorders of consciousness (e.g.,
unresponsive wakefulness state, minimally conscious state, locked in syndrome) and healthy subjects in different sleep stages. In addition, sedation with different anesthetics can be differentiated (e.g., Propofol and Ketamine), whereby in both cases patients are behaviorally unresponsive but in case of Ketamine they report vivid dreams (e.g., Casali et al., 2013; Sarasso et al., 2015; Casarotto et al., 2016; Bodart et al., 2017; Rosanova et al., 2018). Moreover, a recent study on MEG revealed increased lempel-ziv-complexity (LZC) complexity during a psychosomatic state of consciousness induced using Ketamine, LSD, and Psilocybin compared to a placebo effect (Schartner et al., 2017). Regarding other, related, patient populations, schizophrenic patients have higher LZC compared to healthy controls (Fernández et al., 2011), and depressed patients have higher MEG pre-treatment complexity that decreases after 6 months of pharmacological treatment (Méndez et al., 2012).

These findings show that there are many valid approaches to use complexity metrics to quantify consciousness, although to our knowledge there are no studies that investigated complexity on early consciousness. Yet, the exact relation of the different metrics is hard to grasp and it is difficult to define a metric that is most suitable for a specific purpose. Therefore, in the current study, as we use a different type of data than previous studies, we follow an explorative approach and apply numerous different metrics to our MEG data (Table 1). The aim of this approach is to capture different aspects of complexity and compare the metrics regarding their behavior towards this special type of data and their usability when employing them to pursue a novel question. In particular, data from fetal MEG recordings with different gestational ages and additional neonatal recordings are used in this study. The data included in the analysis previously showed auditory event-related responses, which allowed for identification of channels with high brain activity within the sensor space. Such data are used, as it is otherwise difficult to localize clusters of brain activity. The goal of the analysis is to evaluate if complexity metrics are a useful tool for fMEG analysis in the search for fetal consciousness.

**MATERIALS AND METHODS**

**Fetal Magnetoencephalography**

fMEG is a non-invasive tool to measure heart and brain activity in fetuses in the last trimester of pregnancy and in neonates shortly after birth (Preissl et al., 2004). For the recording of fetal- and neonatal data, the SARA (SQUID Array for Reproductive Assessment, VSM MedTech Ltd., Port Coquitlam, BC, Canada) system installed at the fMEG Center at the University of Tuebingen was used (Figure 1). To attenuate magnetic activity from the environment, the device is installed in a magnetically shielded room (Vakuumschmelze, Hanau, Germany). The system includes 156 primary magnetic sensors and 29 reference sensors. The magnetic sensors are distributed over a concave array whose shape is designed to match the maternal abdomen. Based on an ultrasound measurement (Ultrasound Logiq 500MD, GE, UK) prior to the fMEG recording, the position of the fetal head is determined and is marked by a localization coil placed on the maternal abdomen. Three additional localization coils are placed on the spine, left and right side of the subject, to track position changes in relation to the sensor array. An ultrasound directly after the measurement is used to confirm the fetal head position. In case of a change in position, datasets are excluded. Auditory stimulation can be presented via a balloon placed between the maternal abdomen and the sensor array. Neonates get a small, child-appropriate earphone (Ear Muffins from Natus, Biologic, San Carlos, CA, USA), placed on one ear. For neonatal recordings, a cradle is attached to the fMEG device. The neonate is attended by one parent inside the measurement room and is measured asleep or quiet awake.

**Dataset**

For the current analysis, fetal and neonatal data from previously analyzed studies were used (Linder et al., 2014; Morin et al., 2015). Both studies were approved by the local Ethical Committee of the Medical Faculty of the University of Tuebingen (No. 476/2008MPG1 and 339/2010BO1). All participants gave written informed consent in accordance with the Declaration of Helsinki and agreed on reuse of data for additional studies. The auditory stimulation paradigm used in these studies consists of an auditory oddball paradigm with a 500 Hz tone as standard and a 750 Hz tone as deviant. The standard occurs in 80% of times in a pseudorandomized order. Each tone is 500 ms long with an inter-trial interval of 1,500 ms. Forty-five fetal recordings were selected from subjects where auditory event-related responses were detected. They have a gestational age range from 29 to 39 weeks—15 of them in an early phase of the third trimester (29–32 weeks), 15 in the middle (33–36 weeks) and 15 in a later phase (37—39 weeks)—approximately uniformly distributed over the whole age range. Fifteen neonatal recordings were included with an age range from 4 to 46 days (mean = 17.47; SD = 12.68). For all subjects, data with auditory stimulation (“audio”) and data without stimulation (“spont”) is available. As the length of the fetal datasets varies from 6 to 15 min, for all

| TABLE 1 | Overview used metrics. |
|---------|------------------------|
| Concept investigated | Method | Objective |
| Entropy | Multiscale entropy | Measurement of self-similarity of time series by looking at repeating sequences on multiple scales |
| Entropy | Multiscale permutation entropy | Measurement of self-similarity of time series by looking at probability of patterns in data on multiple scales |
| Compressibility | Lempel-Ziv-Welch compression | Quantification of compressibility of time series |
| Fractal dimension | Correlation dimension | Measurement of strangeness of attractor, towards which complex system evolves |
| Fractal dimension | Scale free approaches | Detection of power-law exponent that describes scale free behavior and additionally description of multifractal properties |
of them only the first 6 min were used. The neonatal datasets all have a length of 10 min.

**Data Analysis**

**Preprocessing**

As first step for all datasets, the maternal magnetocardiogram and fetal magnetocardiogram were detected by template matching or using the Hilbert transform algorithm and were subtracted from the relevant signal, through signal space projection (Vrba et al., 2004; McCubbin et al., 2006; Wilson et al., 2008). One of the two methods was selected depending on which method detected the magnetocardiogram more accurately, which is the established procedure for fetal brain analysis (e.g., Linder et al., 2014). Matlab R2016b was used (The MathWorks, Natick, MA, USA) for all further processing steps, except for the calculation of the compressibility measure, where Python 2.7 was used (Python Software Foundation). Fieldtrip (Oostenveld et al., 2011) was used to filtering and downsampling of data. All fetal data were filtered between 0.5 and 10 Hz and all neonatal data between 0.5 and 15 Hz which is the usual filtering range for the analysis of event-related brain responses (Linder et al., 2014; Schleger et al., 2014). Data were downsampled from 610.35 Hz to 256 Hz. An example dataset can be seen in Figure 2.

In the previous analysis of these datasets, five channels were selected for the analysis of auditory event-related responses (example for analysis procedure in Schleger et al., 2014). As these were the five channels with the highest evoked brain activity, they were selected as brain channels for this study (“brain”). Additional to those five channels, five control channels (“control”) were selected which were all more than 15 cm away from the previously selected brain channels. By default, the five channels were situated in the upper middle part of the sensor field. This was suitable for all neonatal recordings due to the similar positioning of the neonate on the sensor array. In fetal recordings, if the fetus was positioned in a way that was too close to the default control channels, resulting in selected brain channels with a distance less than 15 cm, five channels in the lower middle part were selected as control channels. As the five brain channels showed similar behavior, for simplicity of the later analysis only the brain channel with the highest amplitude, and one control channel, were used. Distribution of selected channels within sensor array can be seen in Figure 3.

**Power Spectral Analysis**

To provide a more traditional view on the data, we also calculated power spectral density, using the Fast Fourier Transform algorithm implemented in Matlab. Unlike described in the preprocessing section, data were pre-filtered from 1 Hz to 35 Hz to allow a better display of power in relation to frequency. All other steps were the same. Mean power in the delta range (1.5–3.5 Hz) and theta range (4–8 Hz) was calculated for further comparisons.

**Complexity Metrics**

To measure the informational complexity of the fMEG signal, multiscale entropy (MSE) and multiscale permutation entropy (MPE) were used, as well as LZC. Additionally, we included the geometrical properties of the signal and measured its amount of

---

1 https://www.python.org/
fractality. As an approximation for this, the CD, also known as dimensional complexity (Janjarasjitt et al., 2008b), is calculated. In addition, scale-free behavior, which is a basic property of a fractal, was taken into account.

**Multiscale Entropy**

MSE calculates SE for different time scales. If SE is calculated, lower values indicate more self-similarity in time series and the calculation is largely independent of recording length and relatively consistent (Richman and Moorman, 2000). This traditional entropy measure takes only one scale into account, therefore MSE uses a coarse-grained time series with different scaling factors which takes long range correlations into account (Costa et al., 2002). By considering multiple scales, both highly deterministic and completely random signals result in low values, only complex signals can reach a high value (McIntosh et al., 2008).

For the coarse graining step, the dataset is divided into non-overlapping windows and the data points in each window are averaged. SE is based on the definition of Kolmogorov entropy and is defined as the negative logarithm of the probability of two sequences that are similar for \( m \) points, to be similar at the point \( m + 1 \) as well (Richman and Moorman, 2000). Equation 1 describes the calculation of SE for one scale whereas \( B \) is the number of template matches of length \( m \) and \( A \) the total number of matches of length \( m + 1 \). \( r \) describes the tolerance within which two points are accounted as similar (Richman and Moorman, 2000).
For the calculation of MSE, the “msentropy” function out of the WFTB Toolbox (Goldberger et al., 2000; Silva and Moody, 2014) was used. The parameters used in the current analysis are guided by the default parameters described by Costa et al. (2005). We used $m = 2$, $r = 0.15$ (15% of the standard deviation of the time series) and $N = 15360$ which is a bit lower than the recommended $N = 2 \times 10^4$ but was selected in terms of comparability of the different methods and did not show any disadvantage compared to an example analysis with a larger $N$. We chose a window size between 1 and 20 to calculate the SE for 20 different scales whereas scale one equals the original time series. A time series is considered as more complex than another if a majority of scales show higher entropy values (Costa et al., 2005). For that reason, we used the average MSE over all scales for all further comparisons. To get an impression on the behavior of MSE over different scales, values are displayed in Figures 4A,B.

Multiscale Permutation Entropy
MPE calculates permutation entropy (PE) for multiple time scales. In comparison to other complexity measures, PE is very robust towards noise (Bandt and Pompe, 2002; Zanin et al., 2012). PE looks at different patterns within a time series with the idea that those patterns do not have the same probability of occurrence and that this probability can be informative regarding the underlying dynamics of the system (Zanin et al., 2012). PE uses short samples of a time series to look at their permutation patterns and their frequency of occurrence in relation to all possible permutation patterns (Bandt and Pompe, 2002). PE can be used to quantify complexity of a dynamical time series as it refers to its local order structure (Ouyang et al., 2013). A large value of PE indicates that all permutations are equally likely, a value close to zero signifies a very regular time series (Ouyang et al., 2013).

MPE was calculated using the “MPerm” function by Ouyang (2012, November 21). Like for MSE, the first step is a coarse graining where we selected the same time windows as in the MSE calculation. In this study, the short samples had a length of 4 and a time delay of 1 which corresponds to the default values. Like for the MSE, the average MPE over all scales was used for further comparisons. Values of MPE over different scales are displayed in Figures 4C,D.

Lempel-Ziv-Welch Compressibility
LZC is a measure closely related to Kolmogorov complexity and Shannon entropy (Gao et al., 2011), and is originally described by Ziv and Lempel (1978). For LZC, a dictionary that starts with the shortest new sequence in a time series is built and then adds longer sequences until it captures all non-repetitive sequences (Ruffini, 2017b). The length of this dictionary defines the amount of compressibility of a time series (Ziv and Lempel, 1978). LZC values increase with increasing frequency but not with increasing amplitude as well as with increasing power of noise and increasing signal bandwidth (Aboy et al., 2006).

To calculate LZC, a signal of length $n$ has to be binarized (with an alphabet with $A = 2$ symbols), which in our case is done by a median split as a threshold. In particular, values below the median are indicated as zero, whereas values above the median as one. The median split is relatively robust to outliers compared to other methods (Aboy et al., 2006). After the binarization process, the data can be compressed to a set of “words,” $c(n)$, and the description length of the dictionary (ILZC) is defined as the number of included words times the bits needed to encode those words plus the bits.

Multiscale permutation entropy (MPE), version 1.3. Retrieved from https://ch.mathworks.com/matlabcentral/fileexchange/37288-multiscale-permutation-entropy-mpe-
needed to define a new symbol in the dictionary (Equation 2; Ruffini, 2017a).

\[ ILZC = c(n) \log_2 \left[ c(n) + \log_2 A \right] \]  

(2)

The complexity counter, \( c(n) \), is then normalized by the length of the data string (Ruffini, 2017b). In the current analysis, \( \rho 0 \), which represents the \( c(n) \) normalized by the original string length, is the value used to indicate LZC. A higher \( \rho 0 \) value indicates higher complexity, thus, less ability to compress. For a more detailed description of the process of compression, the reader is referred to Aboy et al. (2006), and for the algorithm used in the current analysis to Ruffini (2017b).

**Correlation Dimension**

The CD is a measure for the strangeness of an attractor, which is closely related to the fractal dimension D. In addition to the geometrical properties of the attractor it takes the dynamics of coverage of the attractor into account (Grassberger and Procaccia, 1983b). The CD uses the statistics of pairwise distances to estimate dimension and is based on the scaling of mass with size (Theiler, 1990). Correlations between points of long-time series on the attractor are used for that (Grassberger and Procaccia, 1983b). Stationarity of this time series is a requirement to obtain reliable results (Theiler, 1986). Therefore each time window was tested for stationarity with the Augmented Dickey Fuller-Test (implemented in the Matlab Econometrics Toolbox) and non-stationary time windows excluded from analysis (2% of time windows). For a complete description of a higher dimensional nonlinear system, a time series—which is an observation in one dimension—has to be unfolded into a higher dimensional space, the so called “embedding space” (Janjarasjitt et al., 2008a). The process of embedding a time series in a higher dimension is described by Takens (1981).

The CD is calculated using the correlation integral, defined as the ratio of the distances between any two points that are smaller than a certain radius and all possible distances (Theiler, 1990). To determine the distance in the time series of the points to be correlated, the time delay \( \tau \) is introduced. \( \tau \) can be set with the help of the autocorrelation function (Janjarasjitt et al., 2008a).

For the current analysis, the “gencorint” function was used to determine the CD (Grassberger and Procaccia, 1983a; Albano et al., 1988; Theiler, 1986; out of the Chaotic systems Toolbox,
The self-similarity of a fractal can be expressed by a mathematical power-law with a distinct exponent (Eke et al., 2002). One of these exponents is the Hurst exponent (H), which is also referred to as “self-similarity parameter” (Zilber, 2014). It expresses the probability that an event is followed by a similar event and is related to the fractal dimension (D) by \( D = 2 - H \). A value of \( H = 0.5 \) shows that a time series is uncorrelated (e.g., white noise). \( 0.5 < H \) is an indication for long range correlations and \( H < 0.5 \) for long-range anti-correlations (Kantelhardt et al., 2002). For some fractal processes one power-law exponent is not enough to characterize them and a multifractal formalism can be used to describe several exponents (Di Ieva, 2016). In case of multifractality, the scaling function of a signal is not linear anymore and can thus not be described by a single scaling exponent H but by a nonlinear scaling function (Zilber, 2014). Those multiple H are called Hölder exponents and can be used to span a multifractal spectrum. The width of this spectrum (M) is a measurement for the amount of multifractality (Zilber, 2014).

Self-similar processes also described as 1/f or scale free behavior can be mainly observed in the infraslow frequency range of the power spectrum (Zilber, 2014). This knowledge is important for the selection of scale ranges in scale free analysis. For the following analysis, the focus is on self-similar processes in the range 0.5–2 Hz. To assess H and M, two methods for multifractal analysis are employed. For both, all data were normalized to avoid the influence of amplitude differences.

**Multifractal Detrended Fluctuation Analysis**

The multifractal detrended fluctuation analysis (MF DFA) is a robust analysis for the estimation of the multifractal spectrum of power-law exponents of a natural time series (Ihlen, 2012). Its basis is the detrended fluctuation analysis (DFA), one of the most popular methods to estimate scale-free behavior in physiological signals which follows the idea that fluctuations within a signal are following a power-law as a function of the number of sample points (Zilber, 2014). The root mean square (RMS) of a signal is calculated over different scales with a certain number of points within each segment. Larger scales are more affected by slower fluctuations, smaller scales more by faster fluctuations (Ihlen, 2012). For each scale, the RMS of the individual segment—local fluctuation—is calculated, and an overall RMS is computed from these values. The slope of the values of this overall RMS over different scales equals H. In case of MF DFA this calculation is done for multiple orders. As a result, MF DFA obtains the set of weighted overall RMS values whose slopes obtain Hurst exponents for multiple orders. They can be used, to trace back the multifractal spectrum and its width M (Ihlen, 2012). For a more detailed description of the calculation steps, see Supplementary Material and Ihlen (2012).

The MF DFA toolbox (Ihlen, 2012, MF DFA1 algorithm) was used in this study. As the algorithm is built for random-walk like signals—which are integrals of noise-like signals (Zilber, 2014)—first we checked whether the data resembles noise or random-walk signals. This is determined by the value of H, which was calculated by a simple DFA. As a cutoff \( H = 1.2 \) was selected (< noise like, > random walk like; after Ihlen, 2012). If the signal is noise like, it is transformed into a random-walk signal, and if it is random walk like, this step is skipped. As a second step, a general linear detrending of the signal is performed. For the current analysis 19 equally spaced scales ranging from 128 to 512 data points, within a scale, were chosen. The orders were selected to range from −5 to 5 in steps of 0.1. The value of \( H_q \) at the order 2 equals H (Ihlen, 2012). The scale-free parameters H and M were evaluated in the further analysis.

**Wavelet-Leader Based Multifractal Formalism**

Wavelet-leader based multifractal formalism (WLBMF) poses a fast, theoretically efficient and robust analysis method for multifractal properties of real-world data (Zilber, 2014). It uses wavelet leaders to derive the multifractal properties of a signal by the knowledge of the scaling exponents. This can be done because wavelet leaders precisely reproduce the Hölder exponents of a signal (Wendt et al., 2007). Wavelet leaders are defined as the maximum wavelet coefficients within a predefined segment (Ciuciu et al., 2008). The log-cumulants (c1-c3) of the scaling exponents give information about the shape of the multifractal spectrum. Whereas c1 equals its maximum, c2 its width and c3 its asymmetry (Wendt et al., 2007; Zilber, 2014). For a more detailed analysis of the current dataset, we used wavelet leaders to derive the multifractal properties of a signal by the knowledge of the scaling exponents.

| Method      | Computational effort | Parameter space | Comparability | Overall usability |
|-------------|----------------------|-----------------|---------------|------------------|
| MSE         | 3.11 s               | 3               | high          | OK               |
| MPE         | 5.69 s               | 3               | moderate      | OK               |
| LZC         | 0.29 s               | 1               | high          | good             |
| CD          | 37.8 min             | 7               | moderate      | weak             |
| MF DFA      | 7.09 s               | 6               | low           | weak             |
| WLBMF       | 2.84 s               | 7               | low           | weak             |

Computational effort reflects the time to compute six time windows of one channel of one dataset. Parameter Space quantifies the number of parameters that can be manipulated within the analysis. Comparability relates to the abundance in related human neuroscience literature.

---

3 Chaotic systems toolbox, version 1.0. Retrieved from https://ch.mathworks.com/matlabcentral/fileexchange/1597-chaotic-systems-toolbox
4 Minimum embedding dimension, version 1.0. Retrieved from https://ch.mathworks.com/matlabcentral/fileexchange/37239-minimum-embedding-dimension
description of the calculation steps, the reader is referred to the Supplementary Material.

For the calculation of the WLBMF, the toolbox described in Wendt et al. (2007) was used. A Daubechies wavelet with three vanishing moments was selected as a mother wavelet. The scales for the WLBMF were chosen in accordance with the scales of the MFDFA. To increase the reliability of the results, Wendt et al. (2007) implemented a bootstrapping process to obtain sets of log-cumulants, which opens up new possibilities for statistical testing. In the present analysis, we used 100 bootstraps and then averaged over the bootstrapped values to retrieve the variables of interest. For a more detailed description see Wendt et al. (2007). In the present study, we only evaluated c1 and c2, whereby c1 is supposed to be equivalent to H and c2 to M (Zilber, 2014).

**Statistical Analysis**

For statistical analysis, the results of all time windows of a certain subject and condition were averaged. Those mean values were then tested for normality with a Kolmogorov-Smirnoff test. As they showed to be not normally distributed, groups were compared with a Wilcoxon signed rank test. To determine whether there is a trend over gestational ages, a Pearson correlation was calculated. After Bonferroni correction,
significance levels were set to $\alpha = 0.006$ for the complexity metrics and $\alpha = 0.025$ for the additional power spectral analysis.

RESULTS

Usability of Methods

After testing several methods for determining complexity of fMEG data, a first step was to evaluate those methods for potential future use. As shown in Table 2, we evaluated them in terms of computational costs, parameter space, comparability regarding their prevalence in the literature and summarized this with the term “overall usability.” Calculations were executed on a multiprocessor machine [12 Intel(R) Xeon(R) CPU, X5660, 2.80 GHz; 96 GB RAM] without multicore support. The value for computational cost is the time it took to calculate the results of one subject with one channel and six time windows. LZC was calculated fastest (0.29 s) followed by MSE, MPE and the scale-free metrics (MSE: 3.11 s, WLBMF: 2.84 s, MPE: 5.69 s, MFDEA: 7.09 s). CD took a very long time for calculation (37.8 min). Parameter space is defined as the number of parameters that have to be adjusted within the analysis. Parameter space ranged from 1 (LZC) to 7 (CD and WLBMF). Thereby a low computational effort is favored, as well as a small number of parameters that can be adjusted. Abundance of literature is seen as an advantage, to ensure comparability of results. Prevalence in literature is assessed and subjectively estimated within the literature search for this present work and only encounters related human neuroscience literature. No literature search with formal search criteria was performed. Overall LZC had the best overall usability as it is unequivocal and fast and is therefore preferable for future analysis. Nevertheless, this evaluation concerns the present field and has only limited significance for other research questions and other types of data.

Detection of Brain Activity

Power Spectral Analysis

Figure 5 (top) shows the power spectral density of neonatal, as well as fetal data and the corresponding control channels. For neonatal data, shape of the power spectrum clearly distinguishes between brain and control channel, especially in the range from 0.5 Hz to 15 Hz, which is used for complexity analysis. For fetal datasets, this differentiation is visually less clear, yet, the graph clearly indicates that the major power is in the delta and theta frequency range. Further comparisons in these frequency ranges are shown in Figure 5. Comparison of brain and control channels resulted in highly significant differences for neonatal audio and
spont data in both delta and theta range ($p < 0.001$ in all cases).
For fetal data, a significant difference between brain and control was only detected for audio data in the delta range ($p = 0.001$).
Interestingly, fetal audio and spont data differed in their power in the delta range ($p = 0.006$).

Complexity Metrics

**Neonatal Data**

In the neonatal datasets, the comparison of brain vs. control resulted in a significant difference for both audio and spontaneous data for LZC, MPE and CD (all $p < 0.001$). Differences for MSE in spontaneous data were marginally significant ($p = 0.007$). In all cases, the control channel showed higher values compared to the brain channel. For the scale-free metrics in both cases, H/c1 could not differentiate between brain and control but M/c2 showed a significant difference ($p < 0.001$). In case of M, the channel with the brain activity appeared more multifractal than the control channel, whereas in case of c2, the opposite was observed. For detailed results, see Figures 6, 7.

**Fetal Data**

For the comparison of brain vs. control of in the fetal datasets, no clear tendencies could be found. The LZC calculation, as well as the MPE and CD metric did not reveal a significant difference, neither for audio data nor for spontaneous data. MSE showed a difference for audio data only ($p > 0.001$). Differences between brain and control within the scale-free metrics did not hold after correction for multiple comparisons expect c1 for audio data ($p = 0.006$). For detailed results, see Figures 8, 9. The correlation analysis did not show a trend over gestational age for the fetal brain data for any of the metrics used.

**DISCUSSION**

The usability rating of the different metrics revealed that, concerning the ease of use with fMEG data, LZC was evaluated as best, as it is unequivocal and needs low computational effort. The fractality measures have a high parameter space and therefore forfeit comparability, while entropy measures require a higher computational effort and more parameters to adjust compared to LZC.

In the neonatal population, the channel with brain activity showed lower complexity compared to the control channel, measured by MPE, LZC and CD. As in this scenario, the control channel records environmental noise, it shows us that these
metrics can clearly differentiate between a physiological signal and noise. This differentiation is supported by the results of the power spectral analysis. These rather clear results can be seen as proof of concept for the general usability of complexity metrics in fMEG. The direction of the difference lies in the nature of the comparison as LZC is highest for Gaussian white noise (Aboy et al., 2006) and entropy maximal for random time series (Costa et al., 2005). For fetal data, this comparison produced less clear results—only MSE and one scale-free metric showed a significant difference. This could be due to the fact that the control channel does not consist of pure environmental noise like in neonatal recordings but can also contain leftovers of physiological signals produced by fetus and mother. Magnetocardiographic activity of both, mother and fetus, should be taken into account as major confounds here. Results from power spectral analysis indicated that there is no clear difference between physiological signals and noise like it is seen in the neonatal recordings. This uncertainty makes analysis with metrics, that are not established for these kinds of signals challenging. The different preprocessing steps that need to be performed, to evaluate fetal compared to neonatal data should be considered likewise, as two heart signals need to be detected and removed instead of one. It is worth mentioning in this context, that the magnitude of the maternal heart signal is significantly stronger than the fetal brain signal (Figure 2).

The inconsistency of results from different metrics highlights the challenges of working with complexity metrics as neural correlates of consciousness, as well as the caution one should apply to interpret them. Especially, if a dataset consists of as many different aspects as fMEG data does, the choice of the right metric is crucial. Therefore, there is a need for more systematic, comparative studies, to evaluate the relations of different complexity metrics as well as their sensitivity to small changes in analysis parameters. Entropy measures during sleep, for example, can reverse their direction, depending on the time scale used for calculation (Miskovic et al., 2019). Based on the assumption, that when using multiple scales to calculate entropy, random signals result in lower entropy values than complex signals (McIntosh et al., 2008), in the current study we would have expected higher MSE and MPE values for the brain compared to control channel, especially in neonatal data. Yet, we found the opposite in our results; at least when using a scaling range up to 20. Shapes of MSE curves displayed in Figure 4A indicate that there might be a higher complexity for brain signals in a scaling range larger than 20. As the number
of data points left for SE analysis decreases with increasing scale, testing even larger scales is challenging for the present analysis. Because we are looking at activity in a frequency range below 15 Hz, it is nevertheless possible that larger scales reveal more information about patterns in brain activity compared to smaller scales. This challenge—regarding small changes in analysis parameters—especially accounts for metrics with high parameter space like the scale-free metrics that even showed contradicting results in our analysis (M and c2 in neonatal data). The claim of these measures being equivalent (Zilber, 2014) did therefore not hold for the present results. Besides the analysis parameters, the preprocessing steps play a crucial role. We can see this for example in the relatively low LZC values throughout our results, that are related to the previous bandpass filtering of our data, as LZC values decrease with decreasing signal bandwidth (Aboy et al., 2006). This implies that LZC results from different studies, using different filter settings are difficult to compare.

Further work is necessary to conclusively interpret results from this analysis of fetal MEG recordings. Even if largely explorative, this study shows that complexity metrics can be used for fMEG data and the evaluation gives a guidance for future work. Establishing information based metrics of neural activity for the quantification of consciousness before and shortly after birth still needs additional studies. The broad usage of LZC across a variety of studies, in combination with its use in earlier work on consciousness research, makes this metric especially interesting when pursuing this topic, as results can be compared to other subject populations. Yet, we need a better understanding of each metric and its sensitivity to different aspects of the data as well as their relation to different aspects of complexity, to use these empirical measures of complexity to assess the conscious state of a growing human being. However, as a precise assessment of fetal states is still challenging, the implementation of complexity metrics into fMEG research is a goal, that opens up interesting possibilities.

**ETHICS STATEMENT**

Data used in this work was re-used from two studies previously conducted at our center. Both studies were approved by the local Ethical Committee of the Medical Faculty of the University of Tübingen (No. 476/2008MPG1 and 339/2010BO1). All participants gave written informed consent in accordance with the Declaration of Helsinki and agreed on reuse of data for additional studies.
AUTHOR CONTRIBUTIONS

JM conducted the analysis and wrote the manuscript. SB and EK revised the manuscript. JM, FS and HP conceptualized the study. SB, EK, GR, FW and HP advised the analysis. FS provided preceding analysis. All authors read and approved the final version of the manuscript.

FUNDING

This work was funded by the FET Open Luminous project (H2020 FETOPEN-2014-2015-RIA under agreement No. 686764) as part of the European Union’s Horizon 2020 research and training program 2014–2018. We acknowledge support by the German Research Foundation and Open Access Publishing Fund of the University of Tuebingen.

ACKNOWLEDGMENTS

This work was presented as a poster at the HBP International Conference “Understanding Consciousness: A scientific quest for the 21st century,” June 21st–22nd, 2018 in Barcelona.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnsys.2019.00023/full#supplementary-material

REFERENCES

Aboy, M., Hornero, R., Abásolo, D., and Álvarez, D. (2006). Interpretation of the Lempel-Ziv complexity measure in the context of biomedical signal analysis. IEEE Trans. Biomed. Eng. 53, 2282–2288. doi: 10.1109/tbme.2006.888696

Albano, A. M., Muench, J., Schwartz, C., Mees, A. I., and Rapp, P. E. (1988). Singular-value decomposition and the Grassberger-Proccacia algorithm. Phys. Rev. A Gen. Phys. 38, 3017–3026. doi: 10.1103/physreva.38.3017

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

Baranger, M. (2000). Chaos, Complexity and Entropy. Cambridge: New England Complex Systems Institute.

Bodart, O., Gossieres, O., Wannez, S., Thibaut, A., Annen, J., Boly, M., et al. (2017). Measures of metabolism and complexity in the brain of patients with disorders of consciousness. NeuroImage Clin. 14, 354–362. doi: 10.1016/j.nicl.2017.02.002

Burioka, N., Miyata, M., Cornelissen, G., Halberg, F., Takeshima, T., Kaplan, D. T., et al. (2005). Approximate entropy in the electroencephalogram during wake and sleep. Clin. EEG Neurosci. 36, 21–24. doi: 10.1177/155005940503600016

Casarotto, S., Comanducci, A., Rosanov, M., Boly, M., Sarasso, S., Casali, K. R., et al. (2013). A theoretically based index of consciousness independent of sensory processing and behavior. Sci. Transl. Med. 5:198ra105. doi:10.1126/scitranslmed.3006294

Casarotto, S., Comanducci, A., Rosanov, M., Sarasso, S., Fecchio, M., Napolitani, M., et al. (2016). Stratification of unresponsive patients by an independently validated index of brain complexity. Ann. Neurol. 80, 718–729. doi:10.1002/ana.24779

Cincotti, F., Towle, L., Elbert, T., Ray, W. J., Kowalik, Z. J., Skinner, J. E., Graf, K. E., and Birbaumer, N. (1994). Chaos and physiology: deterministic chaos in excitable cell assemblies. Physiol. Rev. 74, 1–47. doi: 10.1152/physrev.1994.74.1.1

Fernández, A., López-Ibor, M.-I., Turro, A., Santos, J.-M., Morón, M.-D., Hornero, R., et al. (2011). Lempel-Ziv complexity in schizophrenia: a MEG study. Clin. Neurophysiol. 122, 2227–2235. doi: 10.1016/j.clinphys.2011.04.011

Gao, J., Hu, J., and Tung, W.-W. (2011). Complexity measures of brain wave dynamics. Cogn. Neurodyn. 5, 171–182. doi: 10.1007/s11571-011-9151-3

Goldberger, A. L., Amaral, L. A. N., Glass, L., Hausdorff, J. M., Ivanov, P. C., Mark, R. G., 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

Grassberger, P., and Proccacia, I. (1983a). Characterization of strange attractors. Phys. Rev. Lett. 50:346. doi:10.1103/PhysRevLett.50.346

Grassberger, P., and Proccacia, I. (1983b). Measuring the strangeness of strange attractors. Physica D 9, 189–208. doi: 10.1016/0167-2789(83)90298-1

Ihlen, E. A. (2012). Introduction to multifractal detrended fluctuation analysis in matlab. Front. Physiol. 3:141. doi: 10.3389/fphys.2012.00141

Janjarsjıtt, S., Scher, M. S., and Loparo, K. A. (2008a). Nonlinear dynamical analysis of the neonatal EEG time series: the relationship between sleep state and complexity. Clin. Neurophysiol. 119, 822–836. doi: 10.1016/j.clinphys.2007.11.012

Janjarsjıtt, S., Scher, M. S., and Loparo, K. A. (2008b). Nonlinear dynamical analysis of the neonatal EEG time series: the relationship between neurodevelopment and complexity. Clin. Neurophysiol. 119, 1812–1823. doi: 10.1016/j.clinphys.2008.03.024

Kaffashi, F., Scher, M. S., Ludington-Hoe, S. M., and Loparo, K. A. (2013). Analysis of the kangaroo care intervention using neonatal EEG complexity: a preliminary study. Clin. Neurophysiol. 124, 238–246. doi: 10.1016/j.clinph.2012.06.021

Kantelhardt, J. W., Zschiegner, S. A., Koscielny-Bunde, E., Havlin, S., Bunde, A., and Stanley, H. E. (2002). Multifractal detrended fluctuation analysis of nonstationary time series. Physica A Stat. Mech. Appl. 316, 87–114. doi: 10.1016/S0378-4371(02)00383-3

Kennon, M. B., Brown, R., and Di Abarbanel, H. (1992). Determining embedding dimension for phase-space reconstruction using a geometrical construction. Phys. Rev. A 45, 3403–3411. doi: 10.1103/physreva.45.3403

Kuzma, K. (2012). Minimum embedding dimension, version 1.0. Available online at:https://ch.mathworks.com/matlabcentral/fileexchange/57239-minimum-embedding-dimension. Accessed May 14, 2019.

Kostović, I., and Judaš, M. (2010). The development of the subplate and thalamocortical connections in the human foetal brain. Acta Paediatr. 99, 1119–1127. doi: 10.1111/j.1651-2227.2010.01811.x

Lagercrantz, H., and Chaneux, J.-P. (2009). The emergence of human consciousness: from fetal to neonatal life. Pediatr. Res. 65, 255–260. doi: 10.1203/PDR.0b013e3181973b0d
Leontitis, A. (2004). Chaotic systems toolbox, version 1.0. Available online at: https://ch.mathworks.com/matlabcentral/fileexchange/1597-

Moscariello, N., Lancerini, G., Schleger, F., Pulvermüller, F. (1995). Fractal dimension

Mateos, D. M., Erra, R. G., Wennberg, R., and Perez Velazquez, J. L. (2018). Méndez, M. A., Zuluaga, P., Hornero, R., Gómez, C., Escudero, J., Rodríguez-Moreno, E. C., Schleger, F., Preissl, H., Braendle, J., Eswaran, H., Abele, H., et al. (2014). Magnetoencephalographic signatures of numerosity discrimination in fetuses and neonates. Dev. Neuropsycoph. 39, 316–329. doi: 10.1087/8565641.2014.914212
Silva, L., and Moody, G. B. (2014). An open-source toolbox for analysing and processing patient neurophysiological and electroencephalographic databases in Matlab and octave. J. Res. Open. Syst. 2:e27. doi: 10.5334/jors.bi

Tagliazucchi, E., Chialvo, D. R., Sinatra, M., Amico, E., Brichant, J.-F., Bonhomme, V., et al. (2016). Large-scale signatures of unconsciousness are consistent with a departure from critical dynamics. J. R. Soc. Interface 13:20151027. doi: 10.1098/rsif.2015.1027
Takens, F. (1981). “Detecting strange attractors in turbulence,” in Dynamical Systems and Turbulence, Warwick 1980, eds D. Rand and L. S. Young (Berlin, Heidelberg: Springer), 366–381.
Theiler, J. (1986). Spuriously dimension from correlation applied limited to time-series data. Phys. Rev. A Gen. Phys. 34, 2427–2432.

Theiler, J. (1990). Estimating fractal dimension. J. Opt. Soc. Am. A 7, 1055–1073. doi: 10.1364/OSA.7.001055
Tononi, G., and Edelman, G. M. (1998). Consciousness and complexity. Science 282, 1846–1851. doi: 10.1126/science.282.5395.1846

Vrba, J., Robinson, S. E., McCubbin, J., Lowry, C. L., Eswaran, H., Wilson, J. D., et al. (2004). Fetal MEG redistribution by projection operators. IEEE Trans. Biomed. Eng. 51, 1207–1218. doi: 10.1109/TBME.2004.827265

Weiss, B., Clemens, Z., Bódizs, R., Vágó, Z., and Halász, P. (2009). Spatio-temporal analysis of monofractal and multifractal properties of the human sleep EEG. J. Neurosci. Methods 185, 116–124. doi: 10.1016/j.jneumeth.2009.07.027
Wendt, H., Abry, P., and Jaffard, S. (2007). Bootstrap for empirical multifractal analysis. IEEE Signal Process. Mag. 24, 38–48. doi: 10.1109/msp.2007.4286563
Wilson, J. D., Govindan, R. B., Hatton, J. O., Lowry, C. L., and Preissl, H. (2008). Integral approach for fetal QRS detection. IEEE Trans. Biomed. Eng. 55, 2190–2197. doi: 10.1109/TBME.2008.923916
Zanin, M., Zunino, L., Rosso, O. A., and Papo, D. (2012). Permutation entropy and its main biomedical and econophysics applications: a review. Entropy 14, 1553–1577. doi: 10.3390/e14081553
Zanin, M., Ding, H., Liu, Y., Qiu, Z., Zhou, C., Ding, H., and Ye, D. (2009). Neurodevelopment in newborns: a sample entropy analysis of electroencephalogram. Physiol. Meas. 30, 491–504. doi: 10.1088/0967-

Zilber, N. (2014). ERF and Scale-Free Analyses of Source-Reconstructed MEG Brain Signals During a Multisensory Learning Paradigm. (Doctoral dissertation, Université Paris Sud-Paris XI, Paris).
Ziv, J., and Lempel, A. (1978). Compression of individual sequences via variable-rate coding. IEEE Trans. Inf. Theory 24, 530–536. doi: 10.1109/tit.1978.1055934

Conflict of Interest Statement: EK is an employee of Starlab SLU, the company that gave birth to Neuroelectrics in 2011, and GR is a co-founder, shareholder, and employee of Neuroelectrics and Starlab.
The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Moser, Bensaid, Kroupi, Schleger, Wendling, Ruffini and Preißl. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.