Combined head phantom and neural mass model validation of effective connectivity measures

Steven M. Peterson¹ and Daniel P. Ferris²

¹Department of Biomedical Engineering, School of Engineering, University of Michigan, Ann Arbor, MI, USA
²J. Crayton Pruitt Family Department of Biomedical Engineering, University of Florida, Gainesville, FL, USA

Abstract

Objective—Due to its high temporal resolution, electroencephalography (EEG) has become a promising tool for quantifying cortical dynamics and effective connectivity in a mobile setting. While many connectivity estimators are available, the efficacy of these measures has not been rigorously validated in real-world scenarios. The goal of this study was to quantify the accuracy of independent component analysis and multiple connectivity measures on ground-truth connections while exposed real-world volume conduction and head motion.

Approach—We collected high-density EEG from a phantom head with embedded antennae, using neural mass models to generate transiently interconnected signals. The head was mounted upon a motion platform that mimicked recorded human head motion at various walking speeds. We used cross-correlation and signal to noise ratio to determine how well independent component analysis recovered the original antenna signals. For connectivity measures, we computed the average and standard deviation across frequency of each estimated connectivity peak.

Main results—Independent component analysis recovered most antenna signals, as evidenced by cross-correlations primarily above 0.8, and maintained consistent signal to noise ratio values near 10 dB across walking speeds compared to scalp channel data, which had decreased signal to noise ratios of ~2 dB at fast walking speeds. The connectivity measures used were generally able to identify the true interconnections, but some measures were susceptible to spurious high-frequency connections inducing large standard deviations of ~10 Hz.

Significance—Our results indicate that independent component analysis and some connectivity measures can be effective at recovering underlying connections among brain areas. These results highlight the utility of validating EEG processing techniques with a combination of complex signals, phantom head use, and realistic head motion.

Keywords

effective connectivity; independent component analysis; electroencephalography; motion artifact; phantom head

Address for Correspondence: Steven M. Peterson, University of Michigan Biomedical Engineering, 2200 Bonisteel Boulevard, Ann Arbor, MI 48109-2099, stepeter@umich.edu.
1. Introduction

A recent thrust of neuroimaging research has been to measure brain activity during mobile real-world scenarios [1]. Traditional neuroimaging methods, such as functional magnetic resonance imaging (fMRI) and positron emission tomography (PET), require stationary subjects, limiting their use for real-world recordings. In contrast, high-density EEG is a promising method for recording real-world brain dynamics due to its portability and high temporal resolution [2,3]. EEG is affected by low spatial resolution and artifact contamination, making it challenging to extract meaningful cortical information [4]. Blind source separation using independent component analysis can separate out cortical and artefactual sources, reducing the impact of artifact contamination and improving spatial resolution [5,6]. Such high-density, source-localized experiments have been performed during mobile tasks such as treadmill walking, stair stepping, and balance-beam walking [7–9].

While many EEG studies analyze frequency-domain spectral power, understanding the flow of information amongst brain areas using connectivity analysis can provide a more complete understanding of the brain by quantifying interactions amongst cortical regions. However, there are many connectivity measures to choose from. One class of measures originated from Granger causality, which states that one signal influences activity in a second signal if information from the past of the first signal provides information that helps predict the future of the second signal [10]. This idea was developed for 2 signals only, but has since been extended to multichannel data by using multivariate autoregressive modelling [11]. One of these extensions is directed transfer function [12], which was based on the transfer function of the autoregressive model. This was later corrected using normalization to be frequency independent and sensitive only to direct connections, leading to full-frequency directed transfer function (ffDTF) and direct directed transfer function (dDTF), respectively [13]. Another extension of Granger Causality is partial directed coherence, which was based on the model coefficients in the frequency domain [14]. Corrections to make this measure less dependent on scaling and later scale-free have resulted in generalized partial directed coherence (gPDC) and renormalized partial directed coherence (rPDC), respectively [15,16].

In additions to these extensions of Granger causality, there is also Granger-Geweke Causality (GGC) [17]. Additionally, other connectivity measures, such as phase locking value (PLV) and weighted phase lag index (WPLI), do not use multivariate autoregressive models. PLV measures the relative phase between two sources [18], but can include spurious, instantaneous connections due to volume conduction. In contrast, WPLI is based on imaginary coherence, which ignores these instantaneous connections and increases sensitivity to true connections [19]. For our study, we used the debiased WPLI-square estimator from Vinck et al. [19]. Because PLV and WPLI do not act on model-fit EEG activity, they take advantage of averaging results across multiple trials. Due to the limitations of various connectivity measures, many other measures and techniques exist beyond the ones listed here, including methods to identify connectivity using non-stationary burst dynamics that Granger causality measures do not usually account for [20,21].

Due to the abundance of connectivity measures, several studies have attempted to validate these measures. Previous research has used simulated data to generate connection patterns...
with a known ground truth [22]. This has been used to verify connectivity during walking [23] and to show that connectivity measures can be affected by volume conduction [24]. The downside of such modelling is that it usually avoids the non-linearities of the real world, which could potentially violate the assumptions of the measure being validated. Another way to compare measures is by recording EEG from human subjects and then using various metrics for validation [25], but this leads to assumptions about the underlying connectivity pattern in the absence of a ground truth. There is currently a need to validate connectivity measures on real-world ground truth signals.

In addition, there is ongoing debate as to whether connectivity estimation should be performed at the channel or source level [24,26,27]. The concern with source-level connectivity is that it impairs the channel data’s correlative structure, removing important information [28]. On the other hand, source-level connectivity is less influenced by volume conduction and involves specific cortical areas [24]. In addition to the volume conduction, motion artifact becomes a concern in mobile settings [29]. There is also a need to determine if independent component analysis preprocessing can result in accurate connectivity estimation in the presence of real-world volume conduction and motion artifact.

One way to provide a real-world testing with ground truth signals is to use a phantom head with embedded antennae. Head phantoms have long been used in fMRI research to test methods [30] and have also validated EEG source estimation techniques [31,32]. EEG phantom heads mounted upon a moving platform have quantified the effects of head motion and cable sway [33,34]. However, no phantom head studies have validated connectivity estimation techniques. In order to test connectivity measures with complex, multi-frequency signals, we used a neural mass model. Neural mass models are based on the oscillatory properties of neuronal networks, which can be used to generate oscillations at various physiological frequency ranges [35]. By summing the results of multiple neural mass models, complex waveforms can be created [36]. Additionally, interconnections can be created between neural mass model signals to analyze EEG connectivity [23,37,38]. However, these previous studies involved computer simulations and did not account for the real-world effects as a phantom head might.

The purpose of this study was to determine the efficacy of independent component analysis and connectivity measures on signals of varying complexity exposed to real-world volume conduction and motion artifact. We hypothesized that independent component analysis would be able to recover the antennae sources and separate out motion artifact as measured by signal to noise ratio and cross-correlation of the resulting independent components. In addition, we hypothesized that the connectivity estimation measures we used would be able to find the true causal interactions based on the peak frequency for each measure.

2. Materials & Methods

Phantom head setup and antenna signals

Our phantom head consisted of a mannequin head with 8 exposed wire pairs, around which we used a combination of dental plaster, sodium propionate, and water to simulate realistic tissue conductance. See Oliveira et al. for a more complete description [33]. We sent
predefined signals into each antenna using an input/output interface (MicroLabBox, dSPACE GmbH, Paderborn, Germany). We used 6 of the 8 antennae due to memory constraints (Figure 1). 3 antenna signals contained intermittent connections, while the other 3 were distractor signals to see how well independent component analysis and connectivity measures performed in the presence of other signals.

The 3 non-distractor signals were classified as low, mid, and high based on the main frequency component of the signal, as shown in Figure 2. The low signal’s peak frequency was at 6.5 Hz, corresponding to the EEG theta band. Peak frequency for the mid signal was at 10 Hz, corresponding to the EEG alpha band. For the high signal, the peak frequency was at 41 Hz, which corresponded to the EEG gamma band. The experiment included 3 conditions with different antenna signals: 1) signals with a single peak frequency, 2) signals with a smeared, single peak frequency, and 3) signals with two frequency peaks.

We generated complex and physiologically-relevant signals for each antenna using a neural mass model based on previous research [35,38]. We used the neural mass model to generate 6 separate sources with peak frequencies in different EEG bands (delta, theta, alpha, beta, low gamma, and high gamma). These sources were summed together to create each final antenna signal, using the different weightings shown in Table 1.

For each condition, we induced periodic connections between the 3 antenna signals of interest, using the pattern shown in Figure 3. The 6 pre-defined signals lasted for 20 minutes total for each signal condition, with intermittent connections every 2 seconds. We recorded 20 minutes of 128-channel EEG (BioSemi Active II, BioSemi, Amsterdam, NL) from these signals sent through the phantom head, resulting in 100 trials for each type of periodic connection.

It should be noted that for the single dominant frequency condition, we found that the frontmost low-frequency distractor signal (peak frequency of 4 Hz) resulted in suboptimal independent component analysis decomposition. When we analyzed the correlation for each antenna signal between 12 sec trials, we found a much higher autocorrelation for this low-frequency distractor signal (0.42) than any of the other signals, including the low signal (0.01). This suggests that independent component analysis performs better with some data variability. We used a 3.25–4.75 Hz notch filter to remove this signal only for the single dominant frequency condition, which improved the independent component analysis decomposition.

We collected real-world human head motion during gait from one young healthy subject (male), using an inertial measurement unit (APDM, Portland, OR) strapped to his forehead. This subject provided written informed consent, and our protocol was approved by the University of Michigan Health Sciences and Behavioral Sciences Institutional Review Board for the protection of human subjects. We recorded 20-minute conditions each of standing and walking at 0.5 m/s, 1.0 m/s, 1.5 m/s, and 2.0 m/s. This data was converted into trajectories that were replicated by our Notus hexapod (Symétrie, Nimes, FR), similar to a previous study [34]. By mounting the phantom head on top of the hexapod, we could simulate realistic human motion during phantom head recordings. Due to electromagnetic
noise from the hexapod motors when using the MicroLabBox, all EEG recordings with the antenna signals turned on were performed with the motors off (signal data). On the same testing day, we separately recorded EEG motion data with the hexapod turned on and the MicroLabBox disconnected (motion data). We then added the motion data to the signal data during post-processing to approximate EEG signal recordings with motion artifact.

**EEG analysis**

EEG data were processed in EEGLAB using custom Matlab scripts [39]. We high-pass filtered the data at 1 Hz to remove baseline drift. We performed bad channel rejection by identifying channels with notably large standard deviation, a kurtosis above 5 standard deviations, or with uncorrelated activity for more than 1% of the trial time [8]. No channels matched these criteria, so all channels were retained. We referenced the data to the common channel average. For the motion-only data, we performed a fast Fourier transform using Welch’s method to characterize the frequency content of motion artifact at different walking speeds. We then added the motion and signal data, creating separate motion trials for each signal condition. In addition, we added simulated pink noise to maintain a similar 1/f power result to standard EEG studies [40]. We also increased the 60 Hz noise by adding uniform random noise that was bandpass filtered between 59–61 Hz. We then re-referenced to the common average across channels and ran adaptive mixture independent component analysis (AMICA) [41,42], using principal component analysis reduction to 60 components beforehand.

After running independent component analysis, we used the maximum cross-correlation between independent components and the original antenna signals to identify the component associated with each antenna. It is important to note that the sign of each independent component time series can be arbitrary based on the component weights, leading to inverted component data compared to the original signals [43]. This is less of an issue for single-frequency sinusoidal signals, where lagging the component signal can remove the inversion effect. Because we dealt with complex, non-periodic signals, we selected the maximum cross-correlation between each the inverted and non-inverted component time series. We also calculated the power spectra of the 3 antenna signals of interest (low, mid, high) and their corresponding components in order to quantify similarity in frequency content. We then computed scalp maps for each component to visually determine the spatial similarity between the antenna and its corresponding component. In addition to cross-correlation, we calculated signal to noise ratio by using the independent component analysis weights that map channels to components and applying them separately to the signal data and to the combined motion and pink noise data because we collected the signal and noise data separately. Signal to noise ratio was calculated as the mean square of the signal data divided by the mean square of the noise data, converted to decibels.

Connectivity was performed using the Source Information Flow Toolbox (SIFT) [44]. We retained the 5 components that best aligned with the 5 antenna signals used (excluding the front-most distractor components for consistency across conditions). These 5 components were processed in SIFT, using a 500 ms sliding window and 25 ms step size. Each window was detrended. Each connection type had 100 trials per condition. We fit separate
multivariate autoregressive models to our data for each motion and signal condition (and connection type), using Hannan-Quinn information criterion to determine the optimal model order [11]. After fitting and validating the model, connectivity was estimated using dDTF, fDTF, gPDC, rPDC, GGC, WPLI, PLV (Figure 4). We also performed phase-randomized surrogate statistics to determine significantly nonzero connectivity estimates [45]. This uses the same model fitting and connectivity estimation techniques, but applied to phase-randomized data, creating a null distribution. Non-significant values were set to 0.

To reduce the dimensionality of our connectivity data, we averaged the resulting time-frequency connectivity values across the first second of connectivity onset, as shown in Figure 5. We then normalized the averaged results to the maximum value for each condition, allowing comparisons across different measures. We plotted this averaged, normalized connectivity together for all connectivity measures during the stationary condition. To quantify relative accuracy and precision, we computed an average frequency and standard deviation during the stationary condition, weighted by the connectivity strengths at each frequency. We also used the maximum value across frequency bins for each connection to determine how strong each estimated connection was. We compared the stationary condition to the motion conditions using correlation of the time-averaged, significance-masked connectivity. This included a comparison between the connectivity results from the stationary condition and from the signals that were sent into each antenna, which helped determine the effect of the phantom head. While we did use surrogate statistics, we were unable to use statistics to compare across conditions and measures because there was only one subject. We feel that it is reasonable to not have statistics because we have a ground truth to compare to. Similar studies attempting to validate EEG processing and connectivity measures have also not used statistics [25,33].

3. Results

The EEG motion artifact noise from head motion during walking was concentrated at frequencies below 4 Hz (Figure 6). Each walking speed contained different frequency peaks. As walking speed increased, EEG noise data power peaks increased in power and shifted towards higher frequencies. This can be seen in the raw data traces, where faster speeds have larger peak amplitudes and faster oscillatory behavior. At faster walking speeds of 1.5–2.0 m/s, large harmonic frequency peaks can be seen near 2 and 3 Hz.

Independent component analysis performed well in finding the 3 signals of interest in each condition (Figures 7 and 8). The only exception was the low signal during the double peak condition, which was not well-recovered based on the difference in power spectra and low signal to noise ratio. Otherwise, independent components had high signal to noise ratio values ~10 dB or higher that remained consistent at fast movement speeds. In contrast, the signal to noise ratio of the Cz channel started near 10 dB during the stationary condition, but decreased to ~2 dB at the fastest walking speed. Visual inspection of the independent component power and original antenna signal power spectra indicated that volume conduction, head motion, and pink noise mostly added power to the delta (1–4 Hz) and gamma (>30 Hz) power bands. Cross-correlation was above 0.9 for the single peak condition, above 0.8 for the mid and high signals for the other conditions. Based on the
decreased signal to noise ratios and cross-correlations for the low signals compared to the other signals in the smeared peak and double peak conditions, independent component analysis seemed to have the greatest difficulty recovering low-frequency signals.

Autoregressive model validation prior to connectivity estimation showed reasonable model fits to the data, as shown in Table 2. All models across signal and motion conditions had low parameter to datapoint ratios (<0.1), indicating that overfitting was unlikely. Interestingly, the model orders increased slightly for the single peak condition compared to the other two conditions. For all conditions, the likelihood of the residuals being white and consistency values were below the desired levels of 0.95 and 85%, respectively, which likely indicates extra data structure not captured by the model. The negative stability index across all conditions indicated that all models were stable. Overall, the models were stable and appeared to avoid overfitting, indicating that they fit the data well. In addition, we used the same model fit across different connectivity measures, so the fit of each model should not have impacted inter-measure connectivity differences.

Time-averaged estimated connectivity varied among different connectivity measures for the stationary motion condition (Figure 9), with some measures containing frequent spurious results. Most connectivity measures were able to determine the mid signal to high signal connection, validating the use of such measures for estimating connectivity. However, there were clear differences across measures. Both PLV and WPLI frequently found spurious, high-frequency connections. They also correctly estimated connectivity in the low to mid connection during the double peak condition, even though independent component analysis did not recover the original low signal. Both GGC and gPDC also had spurious high-frequency connections, with GGC identifying no true connectivity during the single peak condition. In addition, rPDC incorrectly estimated spurious low-frequency connectivity. While ffDTF and dDTF appeared to be robust to noise, we noticed that ffDTF can sometimes estimate the connection to be in the wrong direction, such as to both low to mid connections during the smeared peak condition. dDTF may sometimes show directional connections as bidirectional, but it does not appear to indicate incorrect connectivity direction. Still, this suggests caution when interpreting estimated connectivity direction.

The weighted average and standard deviation of the time-averaged connectivity highlights differences in accuracy and precision across measures (Table 3). PLV, WPLI, and gPDC had consistently high average frequency and large standard deviations, reflecting their susceptibility to spurious high-frequency connectivity. GGC performed best when estimating the mid to high connections for the smeared peak and double peak conditions. Otherwise, it did not estimate many other connections, indicating that its performance can vary considerably based on experimental conditions and the underlying connections present. ffDTF performs well for the single peak condition, but did not find anything for the low to mid connections during the smeared peak condition. Both dDTF and rPDC appear to perform well, but rPDC appears biased towards low frequencies of 4–5 Hz during some mid to high connections when the low to mid connection is also present.

Correlation between the stationary condition and motion conditions show a complex effect of motion on connectivity estimation (Figure 10). All measures were affected by motion,
which may rely on the quality of the independent component decomposition for each condition. dDTF usually had high correlations close to 1 across motion conditions, except for the double connections for the single peak and smeared peak conditions where correlation dropped almost to 0. rPDC also had consistently high correlations near 1 across motion conditions, except during the smeared peak condition. Both WPLI and PLV only displayed correlations near 1 for the single peak condition, indicating that they may be more susceptible to motion effects for complex signals. It is important to note that these results must be interpreted in conjunction with the stationary condition results. For example, GGC was quite consistent for the single peak condition, but the stationary results show that GGC was consistently finding hardly any connectivity. In addition, the estimated connectivity using the raw antenna signals was consistently different from the stationary condition for all measures, highlighting the effect of real-world phantom head testing.

4. Discussion

We used a novel combination of complex neural mass model signals and a phantom head to validate independent component analysis and connectivity measures under realistic head motions. We found that independent component analysis primarily recovered the original signals of interest and separated out motion artifact. For connectivity estimation, we found variable results across measures and conditions, with most measures able to correctly estimate the underlying connectivity. Measures applied directly to the data, instead of a fitted model, were susceptible to spurious high-frequency connections. In general, dDTF, fDfDTF, and rPDC performed best for our experiments out of the measures we used.

4.1 Motion artifact and independent component analysis

The effect of walking on EEG signals occurred mostly at low frequencies, indicating that slow walking speeds minimally affect EEG results, especially at most physiological frequencies. This has been indicated by other studies [46,47], but differences in cable sway across experimental setups can affect results [34]. We bundled the cables together for this study, which likely decreased the effect of motion artifact. As walking speed increased, the spectral power peaks of the noise data and the frequencies of these peaks increased. This highlights the challenge of computationally removing motion artifact during fast walking and running, where the motion artifact is large and can overlap with EEG frequencies of interest. Dual-layer EEG systems that can subtract out motion artifact appear to be a promising method to mitigate this issue [48].

Independent component analysis performed well in mostly recovering the original signals. We expected this given the frequent use of independent component analysis in EEG research and its ability to recover single-frequency, sinusoidal signals during similar phantom head validation [33]. The consistent signal to noise ratio across motion speeds emphasizes the importance of using blind source separation to minimize the effects of motion, which is why such methods are used often during mobile tasks [6,49]. The cross-correlation results aligned well with the signal to noise ratio results of the recovered independent components of interest. Independent component analysis did not recover the low signal as well as the other signals in the smeared peak and double peak conditions, likely because of the added
pink noise, not motion artifact. The effect of pink noise can be seen by comparing the recovered components’ power spectra to the original antenna signals’ power spectra. Across all signals and movement conditions, low frequency power is consistently increased in the recovered components, which likely leads to the decreased signal to noise ratio seen in the recovered low frequency components. Although motion artifact may be an important concern when analyzing low-frequency EEG activity, our results showed robust motion separation using independent component analysis.

4.2 Connectivity estimation measures

Connectivity measures generally identified the true connections, especially the mid to high connection. This validates the use of independent component analysis and such measures for mobile EEG settings. However, there were substantial differences in performance across measures, especially with regards to finding false positives. We especially noticed this for PLV and WPLI, which estimated connectivity directly from the data instead of using a fitted model. Because these measures utilize trial averaging, their performance likely would have increased with more trials. In addition, many other factors could have altered connectivity estimation, such as choice of reference or type of source localization used [25]. Still, multivariate autoregressive modelling may provide a more robust framework for connectivity estimation than trial averaging.

Out of the estimation techniques using multivariate autoregressive modelling, we found that dDTF, fDTF, and rPDC appear to provide the most reliable estimates. GGC appeared unreliable, especially in the single peak and smeared peak conditions. Other techniques have been used for GGC besides multivariate autoregressive models [50], indicating that the methods used with GGC should be carefully considered beforehand [51]. We also found that fDTF estimated the true connectivity correctly in most cases, but some of its results would lead researchers to conclude that the connection occurred in the wrong direction. This makes fDTF potentially problematic to use if directionality is of particular interest, such as analyzing the connectivity between the cortex and leg muscles. Directionality accuracy appears improved for dDTF and rPDC, but it still appears important to utilize statistical tests to firmly establish a specific directionality. Our results show that no one measure provides a completely clean picture of the true underlying connectivity, suggesting that using multiple connectivity measures may provide the most robust estimates of underlying connections.

Despite the consistent component signal to noise ratio values, connectivity estimation still was impacted by motion and real-world volume conduction. Correlations varied between motion conditions and the stationary condition for all measures, without a clear indication of one measure being most robust to motion in all cases. In general, rPDC and dDTF appear to be the most stable across walking speeds, despite varying estimates during the smeared peak condition. While GGC, WPLI, and PLV were fairly consistent in some conditions, it is important to note that their stationary connectivity estimations were not ideal results, even if they were maintained for different walking speeds. In addition, we found consistently low correlation between the estimated connectivity during the stationary condition and the estimated connectivity performed on the original signals before being sent through the head. This effect was seen across all measures and conditions, indicating that volume conduction...
and noise from real-world recording at the scalp do consistently affect the resulting connectivity estimation [24].

4.3 Limitations

While we were able to validate connectivity under real-world scenarios, our study was limited to a subset of connectivity measures and motion artifact that did not consistently occur during the event of interest. There are many other available measures to estimate connectivity, including coherence, mutual information, and multivariate phase synchronization [52]. We focused primarily on measures based on Granger causality that were available in the SIFT toolbox [44]. In addition, there are many other source localization techniques, such as the various beamforming methods [53]. We also did not look at motion artifact that consistently overlaps with connectivity onset, which would be applicable to EEG studies during locomotion. Any lingering motion artifact following independent component analysis may have had a notable effect on connectivity if time-locked to an event of interest. For this reason, it is important to consider the potential effects of motion artifact, even at slow walking speeds, if it is time-locked to the event of interest. The influence of motion artifact depends on a variety of factors, including the performance of blind source separation of motion and brain sources, the events of interest, and cable sway [34,47]. In addition, several post-processing motion artifact removal techniques have been proposed, which might have potentially improved our connectivity results across walking speeds [6,54,55]. Future phantom head studies similar to the one presented here would help validate how such methods work on ground-truth signals in a real-world setting, allowing researchers to better determine which method should work best in a given situation.

4.4 Conclusions

We validated that several connectivity measures can accurately estimate true connections between complex signals exposed to real-world volume conduction and head movement via a head phantom. Independent component analysis recovered most of the original signals and appeared to separate out motion artifact. We were able to show that performing connectivity on sources from independent component analysis can find the true connections in a real-world scenario, but no one measure performed optimally in every condition. It may be beneficial to use multiple connectivity measures to increase confidence in the estimated connectivity results. Our technique opens up the ability to use complex, ground-truth signals in a real-world environment to validate EEG methods, improving our understanding of how well common EEG methods truly work.

Acknowledgments

We would like to thank Jean Vettel and Manuel Vindiola at the Army Research Lab for providing the neural mass model code and assisting with its implementation. We would also like to thank Andrew Nordin, Krista Kernodle, and the Human Neuromechanics Lab for their help with this project.

References

[1]. Ladouce S, Donaldson DI, Dudchenko PA and Ietswaart M 2017 Understanding Minds in Real-World Environments: Toward a Mobile Cognition Approach Front. Hum. Neurosci 10
[2]. Gramann K, Ferris DP, Gwin J and Makeig S 2014 Imaging natural cognition in action Int. J. Psychophysiol 91 22–9 [PubMed: 24076470]

[3]. Gramann K, Gwin JT, Ferris DP, Oie K, Jung T-P, Lin C-T, Liao L-D and Makeig S 2011 Cognition in action: imaging brain/body dynamics in mobile humans Rev. Neurosci 22 593–608 [PubMed: 22070621]

[4]. Urigüen JA and Garcia-Zapirain B 2015 EEG artifact removal-state-of-the-art and guidelines J. Neural Eng. 12 031001 [PubMed: 25834104]

[5]. Makeig S, Bell AJ, Jung T-P and Sejnowski TJ 1996 Independent Component Analysis of Electroencephalographic Data Advances in Neural Information Processing Systems 8 145–51

[6]. Gwin JT, Gramann K, Makeig S and Ferris DP 2010 Removal of movement artifact from high-density EEG recorded during walking and running J. Neurophysiol 103 3526–34 [PubMed: 20410364]

[7]. Bradford JC, Lukos JR and Ferris DP 2016 Electrocortical activity distinguishes between uphill and level walking in humans J. Neurophysiol 115 958–66 [PubMed: 26683062]

[8]. Peterson SM, Furuki E and Ferris DP 2018 Effects of virtual reality high heights exposure during beam-walking on physiological stress and cognitive loading PLoS One 13 e0200306 [PubMed: 25834104]

[9]. Luu TP, Brantley JA, Nakagome S, Zhu F and Contreras-Vidal JL 2017 Electrocortical correlates of human level-ground, slope, and stair walking PLoS One 12 e0188500 [PubMed: 29190704]

[10]. Granger CWJ. 1969; Measurement of Linear Dependence and Feedback Between Multiple Time Series. J. Am. Stat. Assoc. 74:304.

[11]. Geweke J. 1982; Measurement of Linear Dependence and Feedback Between Multiple Time Series. J. Am. Stat. Assoc. 77:304.
[25]. Mahjoory K, Nikulin VV, Botrel L, Linkenkaer-Hansen K, Fato MM and Haufe S 2017 Consistency of EEG source localization and connectivity estimates Neuroimage 152 590–601 [PubMed: 28300640]

[26]. Kaminski M and Blinowska KJ 2017 The Influence of Volume Conduction on DTF Estimate and the Problem of Its Mitigation Front. Comput. Neurosci 11 36 [PubMed: 28553220]

[27]. Van de Steen F, Faes L, Karahan E, Songsiiri J, Valdes-Sosa PA and Marinazzo D 2016 Critical Comments on EEG Source Space Dynamical Connectivity Analysis Brain Topogr.

[28]. Kaminski M and Blinowska KJ 2014 Directed Transfer Function is not influenced by volume conduction inexpedient pre-processing should be avoided Front. Comput. Neurosci 8

[29]. Kline JE, Huang HJ, Snyder KL and Ferris DP 2015 Isolating gait-related movement artifacts in electroencephalography during human walking J. Neural Eng. 12 046022 [PubMed: 26083595]

[30]. Kneeland JB, Knowles RJ and Cahill PT 1984 Magnetic resonance imaging systems: optimization in clinical use Radiology 153 473–8 [PubMed: 6484181]

[31]. Baillet S, Riera JJ, Marin G, Mangin JF, Aubert J and Garnero L 2001 Evaluation of inverse methods and head models for EEG source localization using a human skull phantom Phys. Med. Biol 46 77–96 [PubMed: 11197680]

[32]. Chowdhury MEH, Mullinger KJ, Glover P and Bowtell R 2014 Reference layer artefact subtraction (RLAS): a novel method of minimizing EEG artefacts during simultaneous fMRI Neuroimage 84 307–19 [PubMed: 23994127]

[33]. Oliveira AS, Schlind BR, Hairston WD, König P and Ferris DP 2016 Induction and separation of motion artifacts in EEG data using a mobile phantom head device J. Neural Eng. 13 036014 [PubMed: 27137818]

[34]. Symeonidou E-R, Nordin AD, Hairston WD and Ferris DP 2018 Effects of Cable Sway, Electrode Surface Area, and Electrode Mass on Electroencephalography Signal Quality during Motion Sensors 18

[35]. David O and Friston KJ 2003 A neural mass model for MEG/EEG: coupling and neuronal dynamics Neuroimage 20 1743–55 [PubMed: 14642484]

[36]. Ma Z. 2018; Neurophysiological Analysis of the Genesis Mechanism of EEG During the Interictal and Ictal Periods Using a Multiple Neural Masses Model. Int. J. Neural Syst. 28:1750027. [PubMed: 28639499]

[37]. Gordon SM, Franaszczuk PJ, Hairston WD, Vindiola M and McDowell K 2013 Comparing parametric and nonparametric methods for detecting phase synchronization in EEG J. Neurosci. Methods 212 247–58 [PubMed: 23085564]

[38]. Vindiola MM, Vettel JM, Gordon SM, Franaszczuk PJ and McDowell K 2014 Applying EEG phase synchronization measures to non-linearly coupled neural mass models J. Neurosci. Methods 226 1–14 [PubMed: 24485868]

[39]. Delorme A and Makeig S 2004 EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis J. Neurosci. Methods 134 9–21 [PubMed: 15102499]

[40]. Linkenkaer-Hansen K, Nikouline VV, Palva JM and Ilmoniemi RJ 2001 Long-range temporal correlations and scaling behavior in human brain oscillations J. Neurosci 21 1370–7 [PubMed: 11160408]

[41]. Palmer JA, Kreutz-Delgado K and Makeig S 2006 Super-Gaussian Mixture Source Model for ICA Lecture Notes in Computer Science pp 854–61

[42]. Palmer JA, Makeig S, Kreutz-Delgado K and Rao BD 2008 Newton method for the ICA mixture model 2008 IEEE International Conference on Acoustics, Speech and Signal Processing

[43]. Makeig S and Onton J 2009 ERP Features and EEG Dynamics: An ICA Perspective Oxford Handbook of Event-Related Potential Components ed Luck S and Kappenman E (New York: Oxford University Press)

[44]. Delorme A, Mullen T, Kothe C, Akalin Acat Z, Bigdely-Shamlo N, Vankov A and Makeig S 2011 EEGLAB, SIFT, NFT, BCILAB, and ERICA: new tools for advanced EEG processing Comput. Intell. Neurosci 2011 130714 [PubMed: 21687590]

[45]. Theiler J, Eubank S, Longtin A, Galdrikian B and Doyle Farmer J 1992 Testing for nonlinearity in time series: the method of surrogate data Physica D 58 77–94
[46]. Nathan K and Contreras-Vidal JL 2015 Negligible Motion Artifacts in Scalp Electroencephalography (EEG) During Treadmill Walking Front. Hum. Neurosci 9 708 [PubMed: 26793089]

[47]. Snyder KL, Kline JE, Huang HJ and Ferris DP 2015 Independent Component Analysis of Gait-Related Movement Artifact Recorded using EEG Electrodes during Treadmill Walking Front. Hum. Neurosci 9 639 [PubMed: 26648858]

[48]. Nordin A, Hairston W and Ferris D 2018 Dual-electrode motion artifact cancellation for mobile electroencephalography J. Neural Eng. 15 056024 [PubMed: 30074489]

[49]. Sipp AR, Gwin JT, Makeig S and Ferris DP 2013 Loss of balance during balance beam walking elicits a multifocal theta band electrocortical response J. Neurophysiol 110 2050–60 [PubMed: 23926037]

[50]. Dhamala M, Rangarajan G and Ding M 2008 Analyzing information flow in brain networks with nonparametric Granger causality Neuroimage 41 354–62 [PubMed: 18394927]

[51]. Dhamala M, Liang H, Bressler SL and Ding M 2018 Granger-Geweke causality: Estimation and interpretation Neuroimage 175 460–3 [PubMed: 29684646]

[52]. Jalili M, Barzegaran E and Knyazeva MG 2014 Synchronization of EEG: bivariate and multivariate measures IEEE Trans. Neural Syst. Rehabil. Eng 22 212–21 [PubMed: 24216751]

[53]. Jonmohamadi Y, Poudel G, Innes C, Weiss D, Krueger R and Jones R 2014 Comparison of beamformers for EEG source signal reconstruction Biomed. Signal Process. Control 14 175–88

[54]. Oliveira AS, Schlink BR, Hairston WD, König P and Ferris DP 2017 A Channel Rejection Method for Attenuating Motion-Related Artifacts in EEG Recordings during Walking Front. Neurosci 11 225 [PubMed: 28491016]

[55]. Arad E, Bartsch RP, Kantelhardt JW and Plotnik M 2018 Performance-based approach for movement artifact removal from electroencephalographic data recorded during locomotion PLoS One 13 e0197153 [PubMed: 29768471]
Figure 1. Phantom head antennae locations.
A CT scan (left) and diagram (right) of the antennae locations within the phantom head are shown, using an axial view. The low, mid, and high antennae were used to generate the signals of interest that contained intermittent connections. These names are based on the peak frequency content of each antenna, with the low signal containing the lowest peak frequency while the high signal included the highest peak frequency of the 3 signals. In addition, we used 3 distractor signals at the antenna locations marked in yellow, which are numbered for later reference. Two other antennae were not used for this study due to technological constraints.
Figure 2. Antenna signals of interest power spectra.
The power spectra for the 3 signals of interest (low, mid, high) are shown for each condition. Signals in the single peak condition have a single sharp frequency peak, indicating one dominant frequency (the smaller peaks in the mid and high signals are from intermittent connections throughout each condition). The power spectra during the smeared peak condition are less sharp, reflecting a more complex signal. For the double peak condition, each signal had two frequency peaks, which is best exemplified by the high signal power spectra.
Figure 3. Connectivity protocol between 3 antennae of interest.
The pattern for each connectivity trial is shown. Circles indicate the 3 antenna signals of interest, with low/mid/high referring to each signal’s relative peak frequency. Arrows signify when a connection between signals was present, with titles at the top indicating what type of connection was present during each 2 second period. Each trial lasted 12 seconds total. We included 100 trials (20 minutes total) for each motion condition.
**Connectivity measures**

**dDTF:** Direct directed transfer function  

**ffDTF:** Full-frequency directed transfer function  

**gPDC:** Generalized partial directed coherence  

**rPDC:** Renormalized partial directed coherence  

**GGC:** Granger-Geweke causality  

**WPLI:** Weighted phase lag index  

**PLV:** Phase locking value

| Used autoregressive model | No autoregressive model |

Figure 4. Connectivity measure abbreviations.  
The abbreviations for the connectivity measures used in this study are provided for reference. We have used different colors to indicate measures that used an autoregressive model (black) and ones that were applied directly to the independent component data (gray).
Figure 5. Example of time-averaged connectivity.
An example of how the time-averaged connectivity is obtained from the time-frequency connectivity results from SIFT. We averaged the 1 second following connection onset, which is at time 0. This results in a one-dimensional trace that shows the average frequency connectivity that measure found. Using this, we were able to plot connectivity results across all measures of interest on a single plot.
Figure 6. Real walking noise effect on EEG.
The time courses (left) and power spectra (right) of just the head motion artifact recorded with the EEG system are shown. We recorded head motion during 5 different walking speeds, from stationary (0 m/s) to 2.0 m/s, and used a motion platform to play back this head motion while recording EEG from the phantom head. Peak frequency power increases at faster walking speeds, along with each peak shifting towards a higher frequency. This can be seen in the time courses, as the rhythmic motion artifact becomes more pronounced and oscillates quicker as walking speed increases.
Figure 7. Antennae signals and recovered independent components.
The results of each independent component decomposition are shown for the 3 conditions. Power spectra results for each movement speed are displayed along with the power spectra of the original signals sent through each antenna in red. We visually compared the time course for each signal before being sent through the phantom (red) to the reconstructed signal from independent component analysis during the stationary condition (green). Additionally, the channel weightings for each independent component are visualized by the inset scalp maps, which match well with the true locations of the antenna that generated the corresponding signal.
Figure 8. Component signal to noise ratio and cross-correlation.
Plots are shown of recovered component signal to noise ratio and cross-correlation between the component and original antenna signal. Signal to noise ratio remained consistent across walking speeds for the components, while the signal to noise ratio of a representative channel (Cz) is notably affected. Additionally, cross-correlation indicated which independent components best match with their respective original signals. With the exception of the low signal for the double peak condition, independent component analysis appeared to recover the original signals well.
Time-averaged connectivity results are shown for our 8 measures of interest for the stationary motion speed only. Connectivity was tested for significance using phase-randomized surrogate statistics with 200 permutations. Non-significant connectivity results were set to 0, and the resulting connectivity was averaged across the 1 second after connection onset. Red titles indicate true connections, with red frequencies indicating the frequency range of the expected connection. Note that WPLI and PLV are undirected measures, so they show the same result regardless of connectivity direction.

Figure 9. Time-averaged connectivity results.
The correlation between time-averaged connectivity of the stationary motion and all other head motions are shown. We also included connectivity results performed on the original signals that were sent into each antenna, designated as ‘no phantom’. All motion speeds (and original signals) were fit to their own model, time-averaged, and masked using phase-randomized surrogate statistics with 200 permutations each. Both dDTF and rPDC appear to have high correlation across motion for most conditions. In addition, connectivity on the signals before they were sent through the antennae had consistently low correlation to the stationary condition, indicating the importance of using head phantoms for validating connectivity methods.
### Table 1.  
Neural mass model frequency weightings for each antenna signal across conditions.

Values show the relative weighting of each source generated from the neural mass model (column headers), with weights adding up to 1. The single peak condition used one neural mass model source for each antenna, while the smeared peak condition distributed weights to neural mass model sources with nearby peak frequencies. The double peak condition used unequal weightings of only 2 neural mass model sources.

|                | Delta (4 Hz) | Theta (6.5 Hz) | Alpha (10 Hz) | Beta (23 Hz) | Low gamma (41 Hz) | High gamma (47 Hz) |
|----------------|--------------|----------------|---------------|--------------|--------------------|--------------------|
| **Low Signal** |              |                |               |              |                    |                    |
| Single Peak    | 0            | 1              | 0             | 0            | 0                  | 0                  |
| Smeared Peak   | 0.1          | 0.5            | 0.25          | 0.1          | 0.05               | 0                  |
| Double Peak    | 0            | 0.7            | 0             | 0            | 0.3                | 0                  |
| Single Peak    | 0            | 0              | 1             | 0            | 0                  | 0                  |
| **Mid Signal** |              |                |               |              |                    |                    |
| Smeared Peak   | 0.1          | 0.1            | 0.5           | 0.25         | 0.05               | 0                  |
| Double Peak    | 0            | 0              | 0.7           | 0.3          | 0                  | 0                  |
| Single Peak    | 0            | 0              | 0             | 0            | 0.7                | 0                  |
| **High Signal**|              |                |               |              |                    |                    |
| Smeared Peak   | 0.1          | 0.1            | 0.25          | 0.05         | 0.5                | 0                  |
| Double Peak    | 0            | 0.3            | 0             | 0            | 0.7                | 0                  |
| Single Peak    | 1            | 0              | 0             | 0            | 0                  | 0                  |
| **Distractor 1**|              |                |               |              |                    |                    |
| Smeared Peak   | 0.5          | 0.1            | 0             | 0            | 0.2                | 0.2                |
| Double Peak    | 0.7          | 0              | 0             | 0            | 0                  | 0.3                |
| Single Peak    | 0            | 0              | 1             | 0            | 0                  | 0                  |
| **Distractor 2**|              |                |               |              |                    |                    |
| Smeared Peak   | 0.1          | 0              | 0             | 0.5          | 0.2                | 0.2                |
| Double Peak    | 0            | 0              | 0.3           | 0.7          | 0                  | 0                  |
| Single Peak    | 0            | 0              | 0.3           | 0            | 0.7                | 1                  |
| **Distractor 3**|              |                |               |              |                    |                    |
| Smeared Peak   | 0.2          | 0              | 0             | 0.1          | 0.2                | 0.5                |
| Double Peak    | 0.3          | 0              | 0             | 0            | 0                  | 0.7                |
Table 2.
Multivariate autoregressive model validation results.

Mean validation results from the fit models are shown, with standard deviation in parentheses. Optimal model order was determined using the Hannan-Quinn Criterion across all time windows. The models were stable and likely avoided overfitting due to negative stability indices and parameter to datapoint ratios below 0.1, respectively. The likelihood of the residuals being white and the model consistency were slightly lower than desired, indicating that the model may not have completely captured all of the data variance. We used the same model fit across different connectivity measures, which avoids differences in model fit from affecting inter-measure differences.

|                               | Single Peak | Smeared Peak | Double Peak |
|-------------------------------|-------------|--------------|-------------|
| Parameter to datapoint ratio  | 0.04 (0.00) | 0.03 (0.00)  | 0.03 (0.00) |
| Model Order                   | 9.4 (0.9)   | 7.8 (0.1)    | 7.8 (0.2)   |
| Residual whiteness likelihood | 0.82 (0.02) | 0.90 (0.01)  | 0.90 (0.01) |
| Consistency (%)               | 75.7 (1.5)  | 74.6 (4.1)   | 75.9 (3.3)  |
| Stability index               | −0.03 (0.00)| −0.11 (0.02) | −0.10 (0.00)|
Table 3. 
Weighted mean and standard deviation of connectivity results.

For each time-averaged connectivity result with surrogate statistics applied, we computed a weighted mean and standard deviation to quantify the location and spread of the estimated connectivity. This is only shown for
the stationary motion speed. In addition, green shading indicates the maximum value within that time-averaged connectivity result to determine how strong the result was.

### Table 3A: Single Peak weighted mean and standard deviation (Hz)

|          | Low → Mid | Mid → High | Low → Mid | Mid → High |
|----------|-----------|------------|-----------|------------|
| dDTF     | 6.5 (2.7) | 10.0 (3.9) | 5.9 (2.3) | 9.1 (2.1)  |
| fDTF     | 6.8 (4.1) | 10.3 (3.8) | 6.0 (2.8) | 8.5 (2.5)  |
| gPDC     | 11.4 (9.7)| 11.4 (5.4) | 8.2 (7.2) | 10.4 (6.7) |
| rPDC     | 9.6 (7.9) | 10.6 (4.7) | 4.6 (4.3) | 11.0 (5.0) |
| GGC      | 10.4 (9.0)| 39.5 (4.0) | 9.5 (4.6) | 39.6 (3.5) |
| WPLI     | 15.3 (12.7)| 12.9 (7.1)| 15.2 (12.5)| 12.2 (9.2)|
| PLV      | 16.2 (12.9)| 22.5 (17.9)| 14.8 (12.6)| 18.8 (15.4)|

### Table 3B: Smear Peak weighted mean and standard deviation (Hz)

|          | Low → Mid | Mid → High | Low → Mid | Mid → High |
|----------|-----------|------------|-----------|------------|
| dDTF     | 5.1 (3.2) | 16.0 (5.8) | 5.3 (3.4) | 10.0 (5.5) |
| fDTF     | -         | 14.3 (3.4) | -         | 9.5 (4.5)  |
| gPDC     | 12.3 (15.2)| 18.9 (13.6)| 10.2 (12.6)| 11.7 (7.2)|
| rPDC     | 3.9 (5.8) | 13.9 (14.6)| 3.7 (4.8) | 4.0 (2.7)  |
| GGC      | 2.5 (0.4) | 15.0 (4.0) | 3.4 (1.1) | 13.4 (9.4) |
| WPLI     | 15.1 (8.9)| 18.3 (10.6)| 16.2 (9.5)| 15.5 (9.1)|
| PLV      | 20.6 (14.7)| 26.6 (16.5)| 22.4 (15.8)| 28.8 (17.1)|

### Table 3C: Double Peak weighted mean and standard deviation (Hz)

|          | Low → Mid | Mid → High | Low → Mid | Mid → High |
|----------|-----------|------------|-----------|------------|
| dDTF     | 4.1 (2.2) | 14.4 (4.7) | 4.7 (1.8) | 12.7 (4.2) |
| fDTF     | -         | 13.5 (2.6) | -         | 12.5 (3.2) |
| gPDC     | 4.6 (5.2) | 17.9 (13.2)| 5.0 (6.1) | 13.8 (11.8)|
| rPDC     | 3.3 (3.3) | 15.2 (11.1)| 3.4 (4.2) | 4.8 (6.4)  |
| GGC      | -         | 13.3 (4.4) | -         | 14.6 (10.5)|
| WPLI     | 21.3 (10.8)| 21.4 (12.0)| 24.0 (12.9)| 21.3 (12.6)|
| PLV      | 23.3 (12.1)| 35.3 (18.4)| 24.6 (13.5)| 35.2 (16.5)|

**Max Value**

0 1