

Power-Law Dynamics in Neuronal and Behavioral Data Introduce Spurious Correlations

Natalie Schaworonkow,^{1,2*} Duncan A.J. Blythe,^{2,3} Jewgeni Kegeles,¹
Gabriel Curio,^{1,2} and Vadim V. Nikulin^{1,2,4}

¹Department of Neurology, Neurophysics Group, Charité University Medicine, Berlin, Germany

²Bernstein Center for Computational Neuroscience, Berlin, Germany

³Department of Computer Science, Machine Learning Group, Technical University of Berlin, Germany

⁴Centre for Cognition and Decision Making, National Research University Higher School of Economics, Russian Federation

Abstract: Relating behavioral and neuroimaging measures is essential to understanding human brain function. Often, this is achieved by computing a correlation between behavioral measures, e.g., reaction times, and neurophysiological recordings, e.g., prestimulus EEG alpha-power, on a single-trial-basis. This approach treats individual trials as independent measurements and ignores the fact that data are acquired in a temporal order. It has already been shown that behavioral measures as well as neurophysiological recordings display power-law dynamics, which implies that trials are not in fact independent. Critically, computing the correlation coefficient between two measures exhibiting long-range temporal dependencies may introduce spurious correlations, thus leading to erroneous conclusions about the relationship between brain activity and behavioral measures. Here, we address data-analytic pitfalls which may arise when long-range temporal dependencies in neural as well as behavioral measures are ignored. We quantify the influence of temporal dependencies of neural and behavioral measures on the observed correlations through simulations. Results are further supported in analysis of real EEG data recorded in a simple reaction time task, where the aim is to predict the latency of responses on the basis of prestimulus alpha oscillations. We show that it is possible to "predict" reaction times from one subject on the basis of EEG activity recorded in another subject simply owing to the fact that both measures display power-law dynamics. The same is true when correlating EEG activity obtained from different subjects. A surrogate-data procedure is described which correctly tests for the presence of correlation while controlling for the effect of power-law dynamics. *Hum Brain Mapp* 00:000–000, 2015. © 2015 Wiley Periodicals, Inc.

Key words: correlations; significance testing; power-law signals; LRTC; EEG; oscillations

Additional Supporting Information may be found in the online version of this article.

Contract grant sponsor: Bernstein Center for Computational Neuroscience, Project B1 (BCCN 01GQ 1001C), a subsidy granted to the HSE by the Government of the Russian Federation for the implementation of the Global Competitiveness Program (to V.V.N.), German Research Foundation, Research Training Group GRK 1589/1 "Sensory Computation in Neural Systems" (to D.A.J.B.).

*Correspondence to: Natalie Schaworonkow, Department of Neurology, Neurophysics Group, Campus Benjamin Franklin, Charité,

University Medicine, Hindenburgdamm 30, 12203 Berlin, Germany. E-mail: natalie.schaworonkow@bccn-berlin.de

Received for publication 17 December 2014; Revised 6 April 2015; Accepted 13 April 2015.

DOI: 10.1002/hbm.22816

Published online 00 Month 2015 in Wiley Online Library (wileyonlinelibrary.com).

INTRODUCTION

Behavioral measures are essential in the study of neuronal activity. In particular, crucial to understanding neuronal function is to identify relationships between behavior (e.g., reaction times) and brain activity measured via different methods (e.g., LFPs: [Zhang et al., 2008], EEG: [Babiloni et al., 2006; Besserve et al., 2008; Bompas et al., in press; Busch et al., 2009; Dijk et al., 2008; Ergenoglu et al., 2004; Hanslmayr et al., 2007; Linkenkaer-Hansen et al., 2001; Monto et al., 2008], fMRI: [Boly et al., 2007; Coste et al., 2011; Fox et al., 2007; Pessoa et al., 2002]). Typically, after data acquisition, a correlation coefficient between behavioral measures and features arising from the measured neuronal activity is calculated. If a significant correlation is found, a functional dependence is concluded.

When performing a correlation analysis, large empirical correlations may indeed result from the presence of genuine underlying functional relations; however, depending on the experimental setup and data analytic steps taken, there exist innumerable factors which may generate a spuriously high empirical correlation. These may include confounding variables or data preprocessing steps, such as low pass filtering, or as is considered in this article, the temporal structure of the behavioral and/or neuronal measures recorded. The effect of such temporal structure on correlation is well known in the statistical time-series literature [Yule, 1926] but its effect in psychology and neuroscience is worth closer scrutiny.

This situation is especially in need of correction since the temporal structure of behavioral responses recorded in e.g. simple reaction time paradigms, and in neural activity, for example, in the amplitude of bandpass-filtered oscillations (e.g. alpha activity) is particularly pronounced. On the single-trial level, responses have been shown to exhibit substantial variability. Such variability may be explained by the presence of power-law dynamics, i.e. resulting from a process displaying fluctuations over a range of time scales. These power-law dynamics may be formally described equivalently in terms of:

- i. The autocorrelation function of the time series decaying as a function of the lag l with exponent γ : $R(l) \sim 1/l^\gamma$.
- ii. The spectral power is inversely scaling as a function of the frequency f with exponent β : $S(f) \sim \frac{1}{f^\beta}$.
- iii. The fluctuation function $F(L)$, as estimated by Detrended Fluctuation Analysis (DFA), increasing as a function of the window length L as α : $F(L) \sim L^\alpha$. Informally, the fluctuation function denotes the variation of the cumulative sum of the time series at hand when viewed over a fixed number of samples, the window length L . The DFA exponent, α , is often referred to as the *Hurst exponent*.

The relations between the various exponents may be calculated by the Wiener-Khinchin theorem: $\beta = 2\alpha - 1$ and $\gamma = 2 - 2\alpha$. In the following, α will be used.

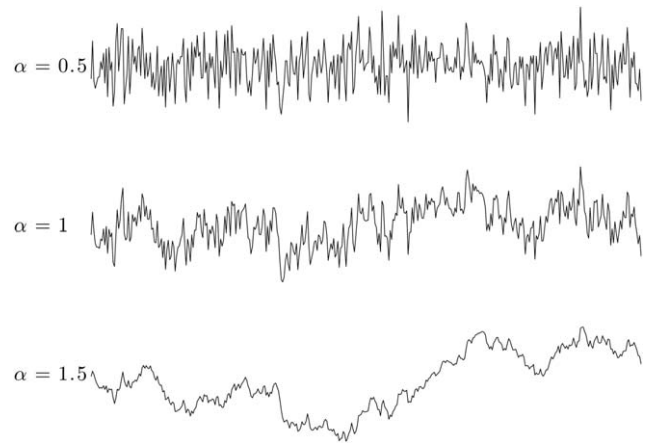


Figure 1.

Visualization of time series with different α -exponents. Top: $\alpha = 0.5$ (white noise), middle: $\alpha = 1$ (pink noise), bottom: $\alpha = 1.5$ (Brownian motion). Smoothness increases with increasing α . The same random seed was used for generation of the three time series.

Time series subject to large α are more predictable in character and smoother in appearance than time series with lower α (see Fig. 1). Time series with $\alpha = 1/2$ may be approximated by white noise in their lower frequencies, time series with $\alpha < 1$ are stationary (for example, fractional Gaussian noise) and with $\alpha > 1$, non-stationary (for example, fractional Brownian motion [Mandelbrot and Van Ness, 1968]).

The presence of power-law dynamics has been established in a variety of behavioral measures (reaction times [Van Orden et al., 2003], memory-related tasks [Maylor et al., 2001; Rhodes and Turvey, 2007], mental rotation [Gilden and Hancock, 2007], and visual discrimination [Gilden and Wilson, 1995]), see Kello et al. [2010] for a review on scaling laws in behavior. Neural activity also exhibits power-law dynamics. This has been established using various imaging methods, including fMRI [Bullmore et al., 2001; Fox et al., 2007] and LFPs [Bedard et al., 2006; Leopold et al., 2003; Milstein et al., 2009]. For M/EEG, power-law dynamics have also been exhibited in raw electrode data [He et al., 2010; Miller et al., 2009; Pritchard, 1992], as well as the amplitude envelopes of band-pass filtered oscillations [Linkenkaer-Hansen et al., 2001; Nikulin and Brismar, 2004, 2005; Palva et al., 2013; Smit et al., 2013]. The simulation results and testing procedures we present here are valid equally for all of the cited paradigms and modalities; as a case study, in this article, we consider neural data in the form of amplitude envelopes of band-pass filtered oscillations. For a recent review on power-law dynamics in brain activity see He [2014].

While the temporal structure of behavioral and neural time-series has been widely studied and found to exhibit long-range temporal correlations (LRTC), otherwise

termed power-law dynamics, the effect in confounding correlation analyses has not been explicitly investigated.

This article provides a critical assessment of correlation analysis between neural activity and behavioral measures in the presence of power-law dynamics in both domains. Significance levels for correlation, resulting from not taking into account power-law dynamics, are shown in simulation and data analysis of multichannel EEG data to be highly erroneous over a range of scaling exponents. We propose a surrogate data procedure for adjusting the significance level for correlation analysis which respects the temporal structure implied by power-law dynamics and which we show to achieve the correct significance level regardless of exponent values.

MATERIALS AND METHODS

Simulations: Generating Time Series Subject to a Given α -Exponent

Independent time series subject to a range of α -exponents were generated using the algorithm proposed by Kasdin [1995]. The time series are generated by filtering white noise processes with an IIR filter, whose coefficients are dependent on the desired α -exponent. The method allows for the generation of time series with a given α -exponent for a wide range of α -values. The length of the time series was chosen to correspond to the number of trials in the EEG experiment (see Data Acquisition section), $n = 800$. The α -exponent was set to a range of values between 0.5 and 1.5 in steps of size 0.005, resulting in 200 time series.

Spearman's rank correlation coefficient ρ was computed between all independent pairs of simulated time series. To obtain an empirical distribution of the correlation coefficient, pairs of time series were generated and a correlation coefficient computed between the pair 1,000 times. The value for which the correlation is assessed as significant was computed (at the significance level $\alpha = 0.05$) by a standard test for correlation, which uses Student's t approximation to the distribution of the Spearman correlation coefficient (we use the implementation in `corr.m` in *MATLAB*). The number of rejections of the null hypothesis according to this significance test was computed for each pair of α values.

To test the effect of the number of trials on significance levels, we further varied the time series length, setting $n = 400, 1,600$, respectively. Removing a trend from RT data is a common preprocessing procedure [Van Orden et al., 2003]. To check for the effect of detrending, we repeated the above analysis after removing a linear trend. Often, a binning procedure is used, combining trials according to percentiles of the measured value, with the aim to arrive at a clearer signal through averaging. To investigate the effect of this procedure, we followed the approach of Linkenkaer-Hansen et al. [2004], with 10 as the number of bins.

Data Acquisition

Recording setup

EEG and EMG data were acquired with BrainAmp MR-plus (Brain Products, Germany) amplifiers and sampled at 1,000 Hz. During acquisition the EEG data were bandpass-filtered in the frequency range of 0.015 to 250 Hz. An electrode cap with 96 electrodes, placed according to the 10 to 20 system, was used. The electrodes were referenced to the nose.

The stimulation electrode was placed on the index finger between the distal and medial phalanges of the non-dominant hand. Bipolar EMG electrodes were placed on the FDI (first dorsal interosseus; reference on second metacarpophalangeal joint) and FDP (flexor digitorum profundus with reference on processus styloideus ulnae) muscles of the nonstimulated hand for the detection of the index finger movement. Additionally, an acceleration sensor was placed on the response finger.

Participants were seated in a comfortable chair. To avoid visual distractions, a dark screen (2 m \times 1.5 m) was placed in front of the participant. Participants were instructed to fixate their gaze on a white fixation cross in the middle of the screen for the duration of the experiment.

Participants and experimental paradigm

Eleven healthy participants with normal or corrected-to-normal vision, with no history of neurological disease, took part in the experiment. Two participants were left-handed, two were female. Participants gave written informed consent and received monetary compensation for their participation. The study was approved by the local Ethics Committee of the Charité University Medicine Berlin, Germany and conformed to the declaration of Helsinki.

Before and after the experiment, the somatosensory threshold of each participant was determined. The experiment employed a simple reaction time paradigm, consisting of nine stimulus blocks, each lasting 5 min. In each block a weak electrical stimulus (square wave current pulse of 0.2 ms duration, three times the determined sensory threshold, so that it was well perceivable) was applied to the index finger of the nondominant hand; the participant was required to react as fast as possible with a light twitch of the index finger of the dominant hand. This setup was fixed for all participants except for one participant, whereby the configuration was flipped (response: nondominant hand, stimulation: dominant hand).

The inter-stimulus interval was chosen to be distributed uniformly in the interval [2.5 s, 3.5 s]. The number of trials acquired amounted to 866 ± 23 trials.

Three rest blocks of five minutes were recorded, in which the participant was not required to perform a task; the rest blocks were spliced randomly between task blocks.

Ten practice trials were conducted before the EEG recording was started.

Data Analysis

Determination of reaction times from EMG

For the determination of reaction time, the onset of movement was determined with the aid of recorded EMG data. The continuous EMG data were high-pass filtered at a cut-off frequency of 10 Hz and the 50 Hz power line noise was removed with a notch filter at 48 to 52 Hz (Butterworth filter, order 4).

After rectification the continuous EMG data were cut into epochs of ranging from -500 ms to $2,000$ ms relative to stimulus onset. The movement onset was determined manually by a trained researcher and consequently verified by another researcher.

Preprocessing and artifacts

The data were analyzed offline using MATLAB (The MathWorks, Inc., Natick). EEG recordings were decimated to a 200 Hz sampling rate. The 50 Hz power line noise was removed with a notch filter at 48 to 52 Hz (Butterworth filter, order 4). Trials with reaction times faster than 50 ms were discarded. Each channel was visually inspected for artifacts with the aid of the variance in all trials in the window $[-2,000, 1,000]$ ms. For each trial, the neuronal activity was visually inspected for the presence of artifacts (muscle and mechanical) with based on the recorded EOG activity, the variance of activity and by the Mahalanobis distance of the maximum power between trials [Nikulin et al., 2008]. Between 1% and 17% of trials were rejected over participants (mean number of rejected trials: 56).

Correlations with laplacian filtered mean amplitude in α -band

The calculation of prestimulus activity and reaction times was performed as follows: data were bandpass-filtered in the EEG alpha-range (8–13 Hz) (Butterworth filter, order 4). A Laplacian spatial filter (subtracting the mean voltage values of the four surrounding electrodes) was applied to each channel, yielding 31 channels. Whenever one or more of the four surrounding electrodes had been rejected, the Laplacian was not computed for that channel and was not taken further into account. (Blythe et al. [2014] studied the effect of spatial filtering on the computation of α -exponents and demonstrated that α -exponents observed in sensor space distort the true values underlying the processes in source space, concluding that spatial filtering, such as Laplacian filtering, is an essential step to the study of Hurst exponents in EEG data.) The amplitude envelope was calculated as the absolute value of the analytic signal, obtained by the Hilbert transform of

the Laplacian filtered signal. For each trial, the activity $-1,000$ ms to -5 ms relative to stimulus onset was extracted and the mean amplitude in this interval was calculated. Spearman's rank correlation coefficient between the mean amplitude for each trial and RT was computed for each channel. For comparison, activity bandpass-filtered in the beta range (15–25 Hz) was analyzed.

As an additional test of the procedure, also Spearman's rank correlation coefficient between amplitude time series for all channels from one subject and amplitudes from all other subjects was calculated.

Quantifying Fluctuations With Detrended Fluctuation Analysis

We use Detrended Fluctuation Analysis [Peng et al., 1994] to quantify the fluctuations in the RT and EEG data, yielding values of the exponent α of a signal $X(t)$. For a comprehensive introduction to DFA see Hardstone et al. [2012]; briefly, a DFA analysis proceeds by computing the fluctuation function $F^2(n)$, which is defined as the variance of the integrated time series ($Z(t) = \sum_{i=1}^t X(t)$) in windows of length n after detrending. Informally, $F^2(n)$ may be viewed as measuring the smoothness in the time-series; summing a highly autocorrelated time-series yields an almost continuous curve, whereas as summing an uncorrelated time-series yields a jagged random walk. The exponent α is defined as the slope of a log-log least squares fit of n versus $F(n)$ ($\sqrt{F^2(n)}$).

For the data analysis we set eight log spaced window lengths $n = 5, \dots, 50$. Note that for DFA analysis each sample relates either to one reaction time or the averaged amplitude of alpha oscillations over 1,000 ms.

Testing for Significant Correlations

As it will be demonstrated in the results section, the presence of high α -exponents leads to high variance in empirical correlations, skewing standard tests for correlation. This may be corrected for if sufficient time-series "of the same type" are made available, yielding the expected range correlations take; in statistics this is known as *the method of surrogate data*. Surrogate time series were generated using the Amplitude Adjusted Fourier Transform (AAFT) algorithm [Theiler et al., 1992]. AAFT serves to generate time-series, which are independent of the measured time-series $X(t)$, but are subject to the same spectrum and distribution, under the assumption that the measured time-series is a monotonic transformation of a Gaussian process.

The steps involved are:

1. A Gaussian white noise time series $Y(t)$ is generated with equal length.
2. $Y(t)$ is reordered so that the ranks of each data point agree with $X(t)$.

3. The phases of the Fourier transform of $Y(t)$ are randomly resampled from a uniform distribution and the inverse transform is computed, to give a resampled Gaussian time-series.
4. $X(t)$ is reordered so that its ranks agree with the resampled $Y(t)$, to give a surrogate time-series.
5. This procedure is iterated to yield multiple surrogate time-series.

The correlation coefficient was then computed for the two original RT and EEG time series and repeatedly between the original RT time series and surrogate time series generated from EEG time series (number of repetitions: 10,000). The P value for the measured correlation is then given by the lowest percentile in which the correlation lies in the correlation coefficient distribution given by the surrogate data. For an illustration of the surrogate data procedure, see supplementary material, Introduction section. MATLAB code for the surrogate data procedure is available at https://github.com/neurophysics/spurious_correlations.

Type I Error Assessed on Reaction Times and α -Amplitude

We may safely assume that reaction times and EEG time series of separate participants are independent, as the interstimulus intervals were randomized for each subject individually. This allows us to assess the number of type I errors when using the standard correlation procedure of Correlations with Laplacian Filtered Mean Amplitude in α -Band section by computing correlations between the RT of one subject and the *alpha*-amplitude of another as well correlations between *alpha*-amplitudes of different participants. For comparison, we also apply the proposed surrogate data procedure (number of repetitions: 1,000). The number of trials was reduced, in each participant, to the minimum number of trials (taking the first N_{\min} trials) across participants so that the time series agreed in length.

RESULTS

Correlations From Simulations

When computing the correlation coefficient between independently simulated power-law time series, the incidence of significant (positive or negative) correlations, as assessed by the t -statistic, increases with increasing α -exponent of the time series (Fig. 2a). This result can be explained when looking at the distribution of correlation values obtained from repeating the simulation 1,000 times (Fig. 2b). While the mean correlation across repetitions is close to zero, as expected from independent signals, the variance of the obtained correlation values increases with increasing α -exponents, i.e. with more slowly decaying autocorrelation. This implies that the probability of achiev-

ing a high (positive or negative) correlation between two independent time series is higher when their autocorrelation decays slowly, i.e. higher exponents increase the variability of Spearman's correlation coefficient.

For a given sample size and a defined significance level, whether or not a correlation is assessed as significant according to standard statistical testing depends only on the magnitude of correlation coefficient. For example, for sample size $n = 800$ and a significance level of 0.05, the correlation is significant for a value greater than $\rho = 0.0693$. ($\rho = T_n / \sqrt{T_n^2 + n - 2}$, $T_n = F^{-1}(p|n-2)$, with F^{-1} being the inverse Student's t CDF, $p = 1 - \frac{0.05}{2}$ (double tail) and degrees of freedom: $n - 2$). This implies that the increase in variance of correlation values will lead to an increase in correlations assessed as significant by the test applied.

As can be seen in Figure 2c, a substantial fraction of correlations reach significance according to the standard test for correlation across a range of α -exponents. The increasing fraction of significant correlations with respect to increasing α -exponent is systematic and not an artifact of a multiple comparisons (otherwise the resulting fraction would randomly fluctuate with respect to the α -exponent). Random fluctuations in the correlation coefficients are expected (as seen in Fig. 2a) over a single iteration of computation of the correlation coefficient for the generated time series. But after repeating the simulation and looking at the fraction of significant correlations reveals a systematic bias in estimation of the desired significance level, using the standard correlation test. For α -exponents between 0.5 and 1, i.e. the range of values typically found in EEG time series, the maximum fraction is as high as 0.6 (excluding identity correlations), implying that up to 60% of significant correlations according to standard testing may constitute spurious detections), i.e. 12 times the chosen significance level.

In addition, neither of the data processing procedures (altering sample size, removing a linear trend or binning) tested solve the problem that with increasing α -exponent, the fraction of significant correlations between independently generated time series greatly exceeds the defined level for type I errors (see Supporting Information for results regarding the additional data processing procedures: Fig. S2 for removing linear trend, Fig. S3 for binning and Fig. S4 for altering the sample size.)

Reaction Time Statistics

Reaction times for the simple reaction time experiment showed a distribution with a long positive tail, see Figure 3 for an example of a representative participant. The bottom part of the figure shows the time course of reaction times. One can identify fast changes and changes on a greater time scale.

The estimated α -exponents for RT time series of all participants are displayed in Figure 4. The values lie between

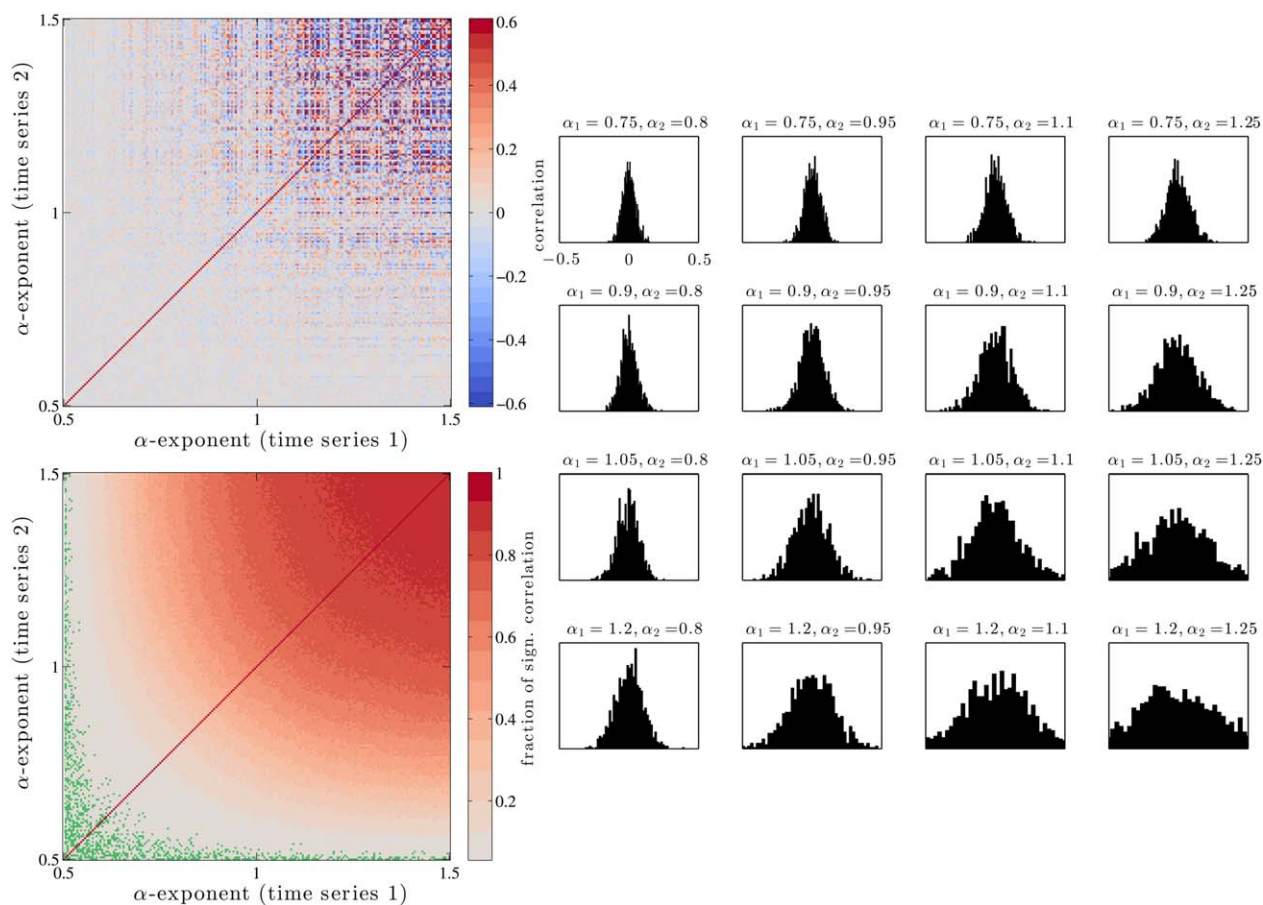


Figure 2.

Top left (a): Simulated pairwise time series correlation (Spearman's ρ) for different combinations of α -exponents. Increasing values of α lead to stronger (positive or negative) correlation values, as evidenced by the more saturated colors towards the upper right region of the graph. Length of time series $n = 800$. Right (b): Distributions of correlation values between simulated time series for changing α -exponent. Bottom Left (c): Fractions of simulated time series with significant correlations ($P < 0.05$)

using standard testing, for different combinations of α -exponents. Fractions that do not exceed 0.05 are plotted in green. This implies that for all combinations of α values plotted in red, standard testing would lead to an increased rate of type I errors. The colormap is discretized for better comparisons. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

0.55 and 0.81, indicating the presence of long-range temporal correlations. ($P = 0.00098$, Wilcoxon signed rank test.)

EEG Time Series Statistics

The estimated α -exponents for Laplacian EEG alpha-band time series are displayed in Figure 4, with the majority of values ranging from 0.5 to 0.8, with a few lying close to 1.4.

Empirical Correlations From Data

Figure 5 displays the RT time course for one subject and the prestimulus alpha-amplitudes for another subject; a

topographic map displaying significant correlations is presented. Despite their independent generation, a large number of significant correlations, according to standard statistical testing procedures, between the RT and EEG data are found. Across all participants, correlating RT with prestimulus alpha-amplitudes yields a significant correlation 535 times of 3,070 possible combinations (17%) (significance level $\alpha = 0.05$). Correlations between prestimulus alpha-amplitudes from different participants are assessed as significant in 15% of cases, using the standard test for correlation. Mean α -exponent in a given pair and magnitude of correlation between prestimulus alpha-amplitudes from different participants were positively correlated ($r = 0.166$, $P < 0.0001$, $n = 85,566$).

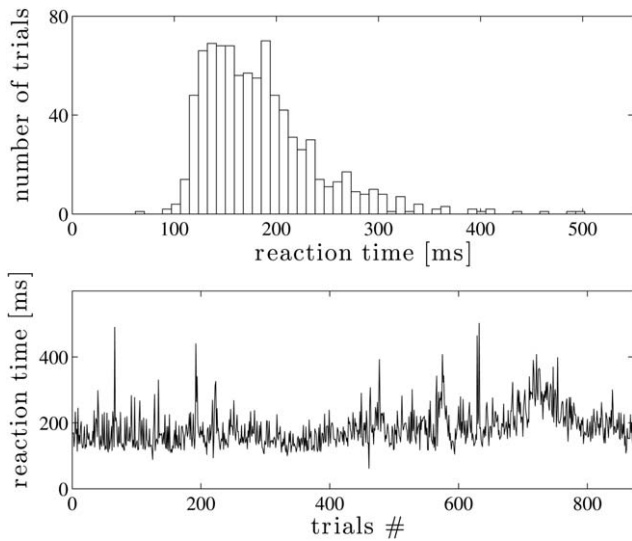


Figure 3.

Exemplary reaction time data from one participant. Top: Histogram of RTs. Bottom: RTs over the course of the experiment, showing 1/f-type fluctuations.

As the number of data points in the case of correlations between RT and EEG data was not sufficient to arrive at estimates of fractions of significant correlation as a function of the α -exponent as calculated in the simulation of Simulations: Generating Time Series Subject to a Given α -Exponent section, surrogate testing was used (in a manner similar to that described in Testing for Signifi-

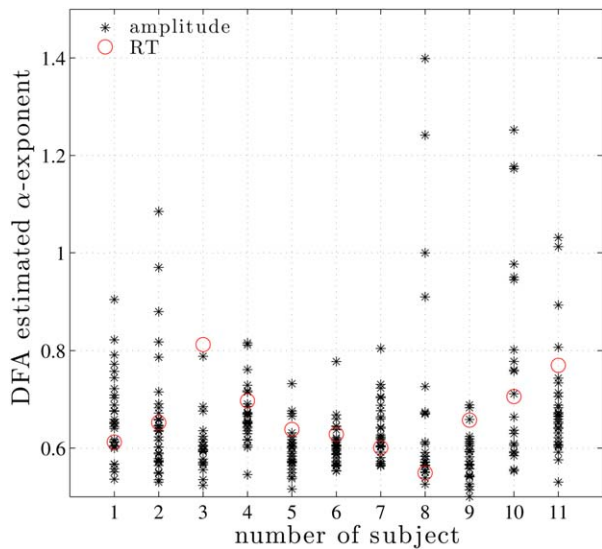


Figure 4.

Estimated α -exponent for RT time series (red circles) and for EEG alpha-band amplitude time series (black stars) for different participants. Each star corresponds to one Laplacian channel. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

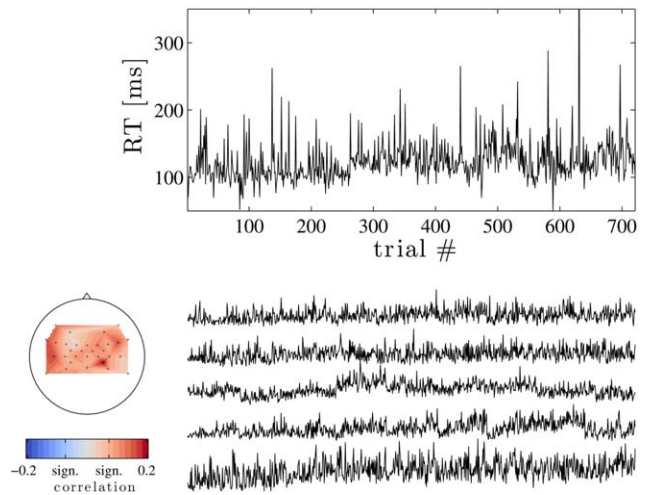


Figure 5.

Top: RT time course for one subject. Below: example EEG alpha-band amplitude time series from five different channels of another subject. Left: Topographical map of correlations between the two types of measures (i.e. RT from one and EEG from another subject). Vertical ticks on the color bar indicate the threshold above which correlations are significant according to parametrical testing. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

cant Correlations section). Surrogate testing, i.e., computing the correlation coefficient between surrogate RT time series from one participant and surrogate EEG time series from all other participants, also yielded a

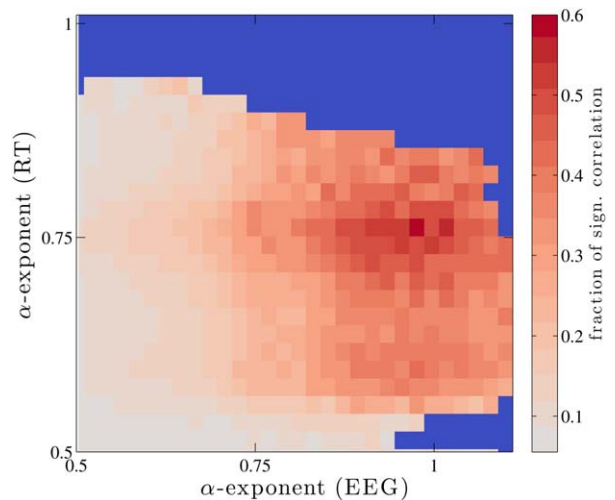


Figure 6.

Fractions of time series with significant correlations ($P < 0.05$) between RT data and Laplace filtered EEG data using standard testing, for different combinations of α -exponents. Combinations of α -exponents that did not occur in the data are plotted in blue. This figure is analogous to Figure 2c. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

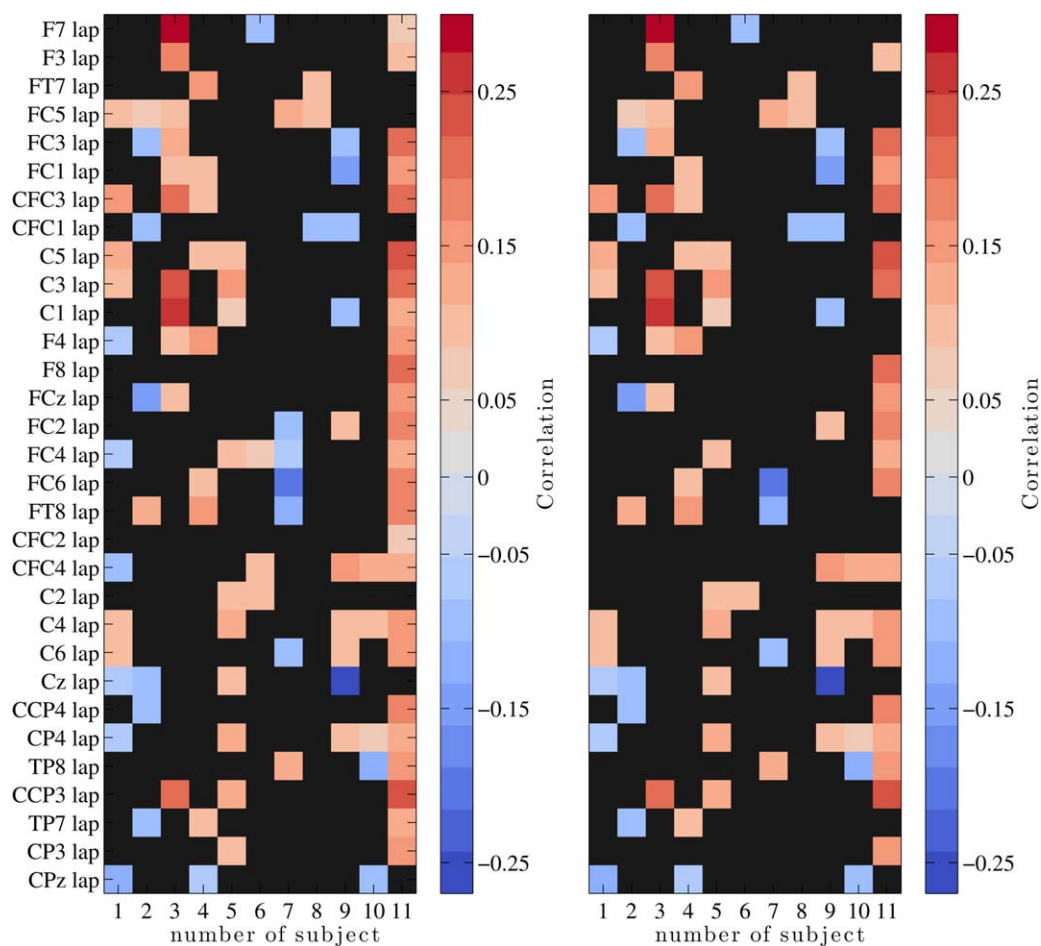


Figure 7.

Correlations between RT and Laplace filtered EEG alpha-band (interval 1,000 to 5 ms before stimulus onset). Correlations that do not reach significance level ($P < 0.05$) are plotted black. Left: uncorrected. Number of significant correlations: 104. Right: significance corrected by the surrogate-data procedure. Number of significant correlations: 89. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

fraction of significant correlations higher than expected according to standard tests for correlation. As we consider here finite length data DFA will return a slightly different (mean standard deviation across participants and channels: 0.03) exponent estimate for each of the 1,000 iterations. To reflect this uncertainty and in order to visualize the irregularly spaced data points, the following procedure was used for plotting: the estimated α -exponents were binned into a grid (two-dimensional histogram). For each bin, the fraction of correlations assessed as significant by the standard test (significance level α evelifi was calculated. Bins with a small number of available data points (< 100) were discarded. Bins for which no data was available are shown in blue.

Figure 6 displays the fraction of significant correlations for Laplace filtered data (1,000 ms prestimulus). There is

an increase in the fraction of significant correlations for higher α -exponents with a maximum fraction of 0.61, implying that as many as 60% of unrelated data (but generated on the basis of real EEG experiments) pairs may display spuriously high correlation.

The results obtained with surrogate data generated from EEG measurements confirm the outcome of the simulation analysis: increasing α -exponents of time series lead to an increased fraction of correlations labeled significant by classical testing. Again, none of the data postprocessing steps examined change the result in a qualitative way, with increasing α -exponents leading to higher fractions of significant correlations using standard testing (see Supporting Information for additional data processing results regarding detrending (Fig. S5) and binning (Fig. S6)).

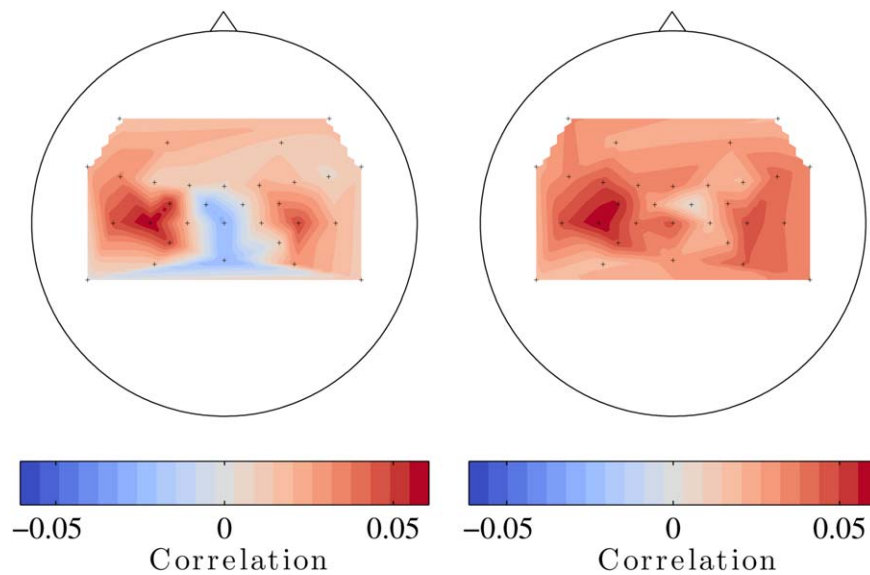


Figure 8.

Topography of correlations between RT and Laplace filtered EEG alpha-band, interval $[-1,000 - 5]$ ms. Left: average of correlations across participants showing larger values over the sensorimotor region. Right: average of absolute values of correlations across participants. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Testing for Correlations: Surrogate-Data Procedure

To investigate the corrected significance levels for correlation between RT and EEG time series from the same participant, which takes the autocorrelation structure of the signals into account, a surrogate-data procedure was used (see Testing for Significant Correlations section for details). Permutation tests are a commonly used method to assess the significance of correlations in place of a parametric test. The question is, what kind of permutation should be used: simply shuffling one time series destroys its autocorrelation structure and yield low correlations when relating the two signals, thus underestimating the variance of distribution of the test statistic, in the presence of autocorrelation.

In this case, the correlation coefficient was computed repeatedly (number of repetitions: 10,000) between the EEG time series with a random phase surrogate (AAFT) RT time series, counting the number of times a correlation at least as large as the original correlation with nonphase-shifted RT time series was achieved. This count divided by the number of repetitions constitutes a corrected P value for evaluating the significance of the correlation.

Figure 7 displays significant correlations for all participants for Laplace filtered mean amplitude in EEG alpha-band 1,000 ms before stimulus onset. The left figure displays all significant correlations, according to the standard test, whereas the right hand figure displays only those confirmed as significant by the surrogate-data procedure.

To illustrate the spatial distribution of correlations assessed as significant, Figure 8 displays the Laplace montage data averaged across participants for Laplace, EEG alpha-band, as well as an average of absolute correlations. As these two topographies display great similarity, it is clear that the direction of correlation (positive or negative) coincides in most cases.

Using the surrogate-data procedure we estimated that fewer correlations are significant than suggested by the standard test for correlation. We see a 14% reduction in significant correlations deemed significant for the EEG alpha-band, 27% for EEG beta-band (results available in supplementary material, see Supporting Information Fig. S7) after application of the surrogate data procedure. Thus application of the surrogate data procedure greatly diminishes the number of correlations assessed as significant.

For an illustration with simulation data that this permutation test preserves the set significance level of 0.05 across the examined range of α -exponents, see the left panel of Figure 9 showing no dependence on the α -exponent for α -exponents smaller than 1, with random fluctuations around a mean value of 0.05. The surrogate data procedure is applicable to pairs of time series where the expected α -exponent are less than 1. This is because the method controls for autocorrelation using the magnitude of the Fourier transform, i.e. the power-spectrum. The empirical power-spectrum is a consistent estimator of the power-spectrum only for $\alpha < 1$ [Kasdin, 1995]. Further work will aim at developing methods which are applicable to time series with $\alpha > 1$. For the applications we are

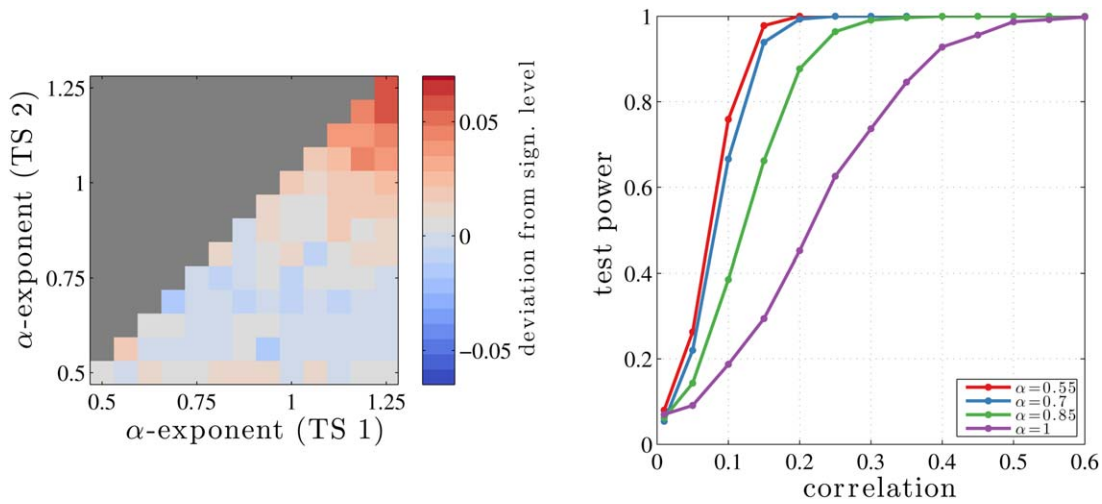


Figure 9.

Left: deviations from the set significance level: empirical fractions of significant correlation using the surrogate data procedure minus the set significance level of $\alpha = 0.05$. Due to computational limitations, the range of α -exponents was limited to [0.5 1.25] in steps of 0.0625, spanning 99.5% of the range of α -exponents present in the EEG data. Right: test power of proposed

surrogate data procedure as a function of the correlation, number of repetitions for each correlation value: 1,000. Due to computational limitations, only time series combinations with the same α -exponent were considered. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

interested in, this is not a severe restriction, since RT time series and amplitude time series of neuronal oscillations are usually subject to an α -exponent in the range from 0.5 to 1 [Linkenkaer-Hansen et al., 2007; Nikulin and Brismar, 2004; Smit et al., 2011]. For an illustration of statistical power, see the right panel of Figure 9. As expected, the test power is dependent on the α -exponent of the signals. To test statistical power of the surrogate-data procedure, the following procedures were performed: two white-noise time series were generated with a defined correlation between them. These time series were then filtered to obtain correlated time series with $1/f$ -spectrum (in accordance to the method described in Simulations: Generating Time Series Subject to a Given α -Exponent section). The significance of the correlation between the two time series was then assessed with the proposed surrogate-data procedure. For each correlation value, 1,000 repetitions of this procedure were performed, calculating the fraction for which the correlation was significant. Due to computational limitations, only time series combinations with the same α -exponent were considered.

aim at investigating the relationship between neuronal activity and behavioral measures through correlation analysis, as both may exhibit power-law dynamics.

Increased Frequency of Spuriously High Correlations With Higher α -Exponents

Power-law dynamics are often present in data and were reported in the range $\alpha \approx 1$ for psychophysical data [Van Orden et al., 2003], for EEG amplitude envelopes [Linkenkaer-Hansen et al., 2001; Nikulin and Brismar, 2004, 2005; Smit et al., 2013], LFP [Milstein et al., 2009] and fMRI [He, 2011; Zarahn et al., 1997]. These values thus fall into a range where the effect demonstrated here arises. The simulations were performed with varying number of trials resulting in up to 60% of correlations assessed as significant, using a standard test, for α -exponents in the range of 0.5 to 1, which corresponds to exponents observed in physiological settings. Thus high correlations do not necessarily point towards neurophysiologically relevant interaction, but may also reflect the univariate signal characteristics of the data.

DISCUSSION: POWER-LAW DYNAMICS CAN RESULT IN SPURIOUS CORRELATION

It has been shown through data analysis and simulations that relating two independent signals with power-law signal characteristics may result in spuriously high correlations. This has implications for a number of lines of research which

It has been shown previously that α -exponents may be altered under certain experimental conditions or in certain experimental groups, for instance in the case of disease (e.g., EEG alpha-band oscillations in Alzheimer’s disease: Montez et al. [2009]), age [Nikulin and Brismar, 2005], vigilance [Bedard et al., 2006], task versus a rest [He, 2011]. The same is applicable to behavioral measures (although

there is less data available explicitly with respect to α -exponents in this case; for instance: RTs for ADHD [Gilden and Hancock, 2007]). Thus differences between groups in the form of significant (as judged by standard methods) correlation for one condition but not the other may be observed simply due to the distinct exponent values of the time series. The present results imply that in order to meaningfully compare correlation values from two conditions, one must report the α -exponents for both conditions.

The effect reported here is independent of the many other problems that may arise in correlation analysis. For instance, the presence of outliers may strongly influence the correlation coefficient, especially Pearson's [Rousselet and Pernet, 2012]. But as simulation results in the previous section showed, using a more robust correlation measure such as Spearman's rank correlation coefficient is still not sufficient to overcome these problems. Another important aspect which recently received attention is proper control for multiple comparisons, which should be performed for instance using a cluster-based approach [Maris and Oostenveld, 2007]. But the spurious correlations described here arise even when only a single test is performed. When performing correlation analysis of time series data, one must therefore take into account the signal properties of the involved time series and examine them for long-range temporal correlations.

Testing for Spurious Correlations

To correct for spurious correlations arising from power-law dynamics a surrogate-data procedure was suggested, with repeated permutations preserving the power-law dynamics present in the time series. Simply shuffling the data would destroy all low frequency components and fail to reveal the influence of power-law dynamics on the correlation coefficient for non-permuted data. When using this surrogate-data procedure to assess the validity of correlations observed between two time series with power-law signal properties, up to one quarter of the correlations labeled as significant by standard testing were not confirmed. To address the question whether RT variability, and behavioral measures in general, can be explained by properties of the neuronal activity preceding the stimulus, tests for spurious correlations should be used to uncover genuine functional relationships. Although there is no widely accepted procedure for testing the correlation between processes with power-law dynamics, the surrogate procedure represents an important step towards general testing methodologies.

Commenting on the neurophysiological implications of the surviving correlations in the simple reaction time experiment is beyond the scope of the article, but a concentration of significant correlations between prestimulus alpha-band amplitude and RT over the sensorimotor cortex of mostly positive sign is in line with previous research [Linkenkaer-Hansen et al., 2004; Zhang et al.,

2008]. As spurious correlations distribute evenly with respect to sign, this trend towards positive correlations may reflect that previously observed correlations are genuine.

Related Work

The effect of autocorrelation structure on correlation and, more broadly, effective sample size has been acknowledged in several lines of research, although not yet explicitly acknowledged in neuroscience. Yule [1926] noted the occurrence of spurious correlations between independent time series as a result of serial correlations between consecutive samples within each variable. Granger and Newbold [1974] dealt with the problem of spurious regressions, demonstrating that regressing two independent random walks resulted in a vast overestimation of a significant relationship between them. Phillips [1986] provided analytical results supporting this argument. Granger went on to develop the cointegration approach [Engle and Granger, 1987], a test whether two time series have a shared drift. This method has not been commonly applied to neuroscientific data.

Analytical derivations for the variance of the correlation coefficient with respect to autocorrelation are given in Bartlett [1935] for AR(1) processes (autoregressive processes of order 1, where the value at one time point depends linearly on one previous time point plus noise), depending on the sample size and autocorrelation parameters: $\text{Var}(C) = \frac{1}{n} \frac{1+\rho_1\rho'_1}{1-\rho_1\rho'_1}$ with ρ_1 being the first order sample autocorrelation of the first time series and ρ'_1 of the second time series. As samples in autocorrelated time series are not independent, this can be used to arrive at an estimated "effective sample size" (the number of independent samples to which the time-series is statistically equivalent) to compute significance. Thiébaux and Zwiers [1984] derived a more general form, including more terms of the autocorrelation. However, they also demonstrated that the effective sample size could not be estimated with good precision, thus rendering such a correction of the significance level unstable.

Surrogate methods have been previously suggested by Theiler et al. [1992], and Ebisuzaki [1997] also suggested a "random phase test" as a permutation test to arrive at a better estimation of significance of correlations between time series in the field of climate research.

Implications

Recent studies [Fox et al., 2005, 2007] have tried to estimate functional connectivity with fMRI by computing the correlation coefficient between raw BOLD signal time series. As these signals also display power-law dynamics, this approach could yield spurious correlations. The significance level is adjusted by calculating the effective number of independent observations. As mentioned

above, estimation of the effective number of independent observations is unreliable and might still overestimate the significance. A recent article [Arbabshirani et al., in press] addressing the effect of autocorrelation on functional connectivity in fMRI finds only small effects of autocorrelations on obtained correlations, when modeling the present autocorrelation with an auto-regressive process of order 1 (AR(1)). This potentially ignores long-term temporal correlations as present in processes with power-law dynamics, leaving the question open whether and how functional connectivity correlations are affected by power-law dynamics.

In studying neuronal oscillations with EEG/MEG, analyzing single trials instead of averages uses a vast amount of additional information about the variability of neuronal processes. However, one has to be aware that features calculated from single trials represent a time series with a temporal structure. The results presented here show that when not taking this into account, the validity of correlations between features such as power or amplitude envelope calculated from EEG and behavior is not ensured. EEG/MEG studies calculating a correlation coefficient not accounting for the temporal structure are numerous.

Estimating a model for the autocorrelation and removing it as in the prewhitening approach [Monti, 2011] additionally has the problem of removing information from the data that is potentially interesting. In the EEG-RT case for instance, the processes underlying slow fluctuations may relate to relatively slow changes in vigilance, while faster fluctuations may relate to the motor preparedness processes, manifested for instance in motor excitability, which is known to be reflected in the amplitude of alpha oscillations [Sauseng et al., 2009]. One goal of research in this field is to find patterns corresponding to both types of processes (slow and fast changes in the amplitude dynamics of oscillations), which is not possible under the prewhitening approach.

EEG studies performing single-trial regression with power as the dependent variable and RT as a regressor [Cohen and Cavanagh, 2011] will also run into the problem of serially correlated residuals. Moving away from single-trial analysis and binning the data [Dijk et al., 2008; Linkenkaer-Hansen et al., 2004] will not necessarily eliminate the effect, as shown in our simulations. There is evidence that sensory detection is also governed by power-law dynamics [Palva et al., 2013], which is of importance for studies considering stimuli close to the sensory threshold, as the behavioral output may result in a time series affected by long-range dependencies.

The problem does not only arise when looking at correlations between brain activity and behavior. As demonstrated, relating EEG amplitudes with each other, leads to an elevated rate of false positives when not controlling for power-law dynamics. An increasing number of studies investigates correlations between brain activity within subject, e.g. looking at correlations between oscillatory proc-

esses between distinct frequency bands [Dähne et al., 2014] or between brain activity from different participants (e.g. “hyper-scanning,” Montague et al. [2002]). If power-law dynamics are not taken into account, spurious correlations can also occur in these paradigms.

In conclusion, power-law dynamics present a problem for evaluating the significance of correlations and may lead to false positives. In the present article, the connection between spurious correlations and power law dynamics was clarified and the dependence of correlation coefficient values on the α -exponent, i.e. the extent of long range correlations of the time series data, was examined. Appropriate methods should be used to ensure valid estimation of significance; a non-parametric permutation test with power-law dynamics preserving surrogate data was proposed to correctly reduce the number of false positives unlike as was observed for parametric methods.

REFERENCES

- Arbabshirani MR, Damaraju E, Phlypo R, Plis S, Allen E, Ma S, Mathalon D, Preda A, Vaidya JG, Adali T, Calhoun VD (2014): Impact of autocorrelation on functional connectivity. *NeuroImage* 102:294–308.
- Babiloni C, Vecchio F, Bultrini A, Luca Romani G, Rossini PM (2006): Pre- and poststimulus alpha rhythms are related to conscious visual perception: A high-resolution EEG study. *Cereb Cortex* 16:1690–1700.
- Bartlett MS (1935): Some aspects of the time-correlation problem in regard to tests of significance. *J R Stat Soc* 98:536–543.
- Bedard C, Kroeger H, Destexhe A (2006): Does the 1/f frequency scaling of brain signals reflect self-organized critical states? *Phys Rev Lett* 97:118102.
- Besserve M, Philippe M, Florence G, Laurent F, Garnero L, Martinerie J (2008): Prediction of performance level during a cognitive task from ongoing EEG oscillatory activities. *Clin Neurophysiol* 119:897–908.
- Blythe DAJ, Haufe S, Müller K-R, Nikulin VV (2014): The effect of linear mixing in the EEG on hurst exponent estimation. *NeuroImage* 99:377–387.
- Boly M, Balteau E, Schnakers C, Degueldre C, Moonen G, Luxen A, Phillips C, Peigneux P, Maquet P, Laureys S (2007): Baseline brain activity fluctuations predict somatosensory perception in humans. *Proc Natl Acad Sci USA* 104:12187–12192.
- Bompas A, Sumner P, Muthumaraswamy SD, Singh KD, Gilchrist ID (2015): The contribution of prestimulus neural oscillatory activity to spontaneous response time variability. *NeuroImage* 107:34–45.
- Bullmore E, Long C, Suckling J, Fadili J, Calvert G, Zelaya F, Carpenter TA, Brammer M (2001): Colored noise and computational inference in neurophysiological (fMRI) time series analysis: Resampling methods in time and wavelet domains. *Hum Brain Mapp* 12:61–78.
- Busch NA, Dubois J, VanRullen R (2009): The phase of ongoing EEG oscillations predicts visual perception. *J Neurosci* 29:7869–7876.

- Cohen MX, Cavanagh JF (2011): Single-trial regression elucidates the role of prefrontal theta oscillations in response conflict. *Front Psychol* 2:30.
- Coste CP, Sadaghiani S, Friston KJ, Kleinschmidt A (2011): Ongoing brain activity fluctuations directly account for inter-trial and indirectly for intersubject variability in stroop task performance. *Cereb Cortex* 21:2612–2619.
- Dähne S, Nikulin VV, Ramírez D, Schreier PJ, Müller K-R, Haufe S (2014): Finding brain oscillations with power dependencies in neuroimaging data. *NeuroImage* 96:334–348.
- Dijk H, van Schoffelen J-M, Oostenveld R, Jensen O (2008): Prestimulus oscillatory activity in the alpha band predicts visual discrimination ability. *J Neurosci* 28:1816–1823.
- Ebisuzaki W (1997): A method to estimate the statistical significance of a correlation when the data are serially correlated. *J Climate Appl Meteorol* 10:2147–2153.
- Engle RF, Granger CWJ (1987): Co-integration and error correction: Representation, estimation, and testing. *Econometrica* 55: 251–276.
- Ergenoglu T, Demiralp T, Bayraktaroglu Z, Ergen M, Beydagi H, Uresin Y (2004): Alpha rhythm of the EEG modulates visual detection performance in humans. *Cognit Brain Res* 20:376–83.
- Fox MD, Snyder AZ, Vincent JL, Corbetta M, Van Essen DC, Raichle ME (2005): The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proc Natl Acad Sci USA* 102:9673–9678.
- Fox MD, Snyder AZ, Vincent JL, Raichle ME (2007): Intrinsic fluctuations within cortical systems account for intertrial variability in human behavior. *Neuron* 56:171–184.
- Gilden DL, Hancock H (2007): Response variability in attention-deficit disorders. *Psychol Sci* 18:796–802.
- Gilden DL, Wilson SG (1995): On the nature of streaks in signal detection. *Cognit Psychol* 28:17–64.
- Granger CWJ, Newbold P (1974): Spurious regressions in econometrics. *J Econometrics* 2:111–120.
- Hanslmayr S, Aslan A, Staudigl T, Klimesch W, Herrmann CS, Bäuml K-H (2007): Prestimulus oscillations predict visual perception performance between and within subjects. *NeuroImage* 37:1465–1473.
- Hardstone R, Poil S-S, Schiavone G, Jansen R, Nikulin VV, Mansvelter HD, Linkenkaer-Hansen K (2012): Detrended fluctuation analysis: A scale-free view on neuronal oscillations. *Front Physiol* 3:450
- He BJ (2011): Scale-free properties of the functional magnetic resonance imaging signal during rest and task. *J Neurosci* 31: 13786–13795.
- He BJ (2014): Scale-free brain activity: Past, present, and future. *Trends Cognit Sci* 18:480–487.
- He BJ, Zempel JM, Snyder AZ, Raichle ME (2010): The temporal structures and functional significance of scale-free brain activity. *Neuron* 66:353–369.
- Kasdin NJ (1995): Discrete simulation of colored noise and stochastic processes and $1/f^\alpha$ power law noise generation. *Proc IEEE* 83:802–827.
- Kello CT, Brown GDA, Ferrer-i-cancho R, Holden JG, Linkenkaer-Hansen K, Rhodes T, Orden GCV, Van Orden GC (2010): Scaling laws in cognitive sciences. *Trends Cognit Sci* 14:223–232.
- Leopold DA, Murayama Y, Logothetis NK (2003): Very slow activity fluctuations in monkey visual cortex: Implications for functional brain imaging. *Cereb Cortex* 13:422–433.
- Linkenkaer-Hansen K, Nikouline VV, Palva JM, Ilmoniemi RJ (2001): Long-range temporal correlations and scaling behavior in human brain oscillations. *J Neurosci* 21:1370–1377.
- Linkenkaer-Hansen K, Nikulin VV, Palva S, Ilmoniemi RJ, Palva JM (2004): Prestimulus oscillations enhance psychophysical performance in humans. *J Neurosci* 24:10186–10190.
- Linkenkaer-Hansen K, Smit DJA, BarkilA, Van Beijsterveldt TEM, Brussaard AB, Boomsma DI, De Geus EJ (2007): Genetic contributions to long-range temporal correlations in ongoing oscillations. *J Neurosci* 27:13882–13889.
- Mandelbrot BB, Van Ness JW (1968): Fractional brownian motions, fractional noises and applications. *SIAM Rev* 10:422–437.
- Maris E, Oostenveld R (2007): Nonparametric statistical testing of EEG- and MEG-data. *J Neurosci Methods* 164:177–190.
- Maylor EA, Chater N, Brown GDA (2001): Scale invariance in the retrieval of retrospective and prospective memories. *Psychon Bull Rev* 8:162–167.
- Miller KJ, Sorensen LB, Ojemann JG, den Nijs M (2009): Power-law scaling in the brain surface electric potential. *PLoS Comput Biol* 5:e1000609.
- Milstein J, Mormann F, Fried I, Koch C (2009): Neuronal shot noise and brownian $1/f^2$ behavior in the local field potential. *PLoS One* 4:e4338
- Montague PR, Berns GS, Cohen JD, McClure SM, Pagnoni G, Dhamala M, Wiest MC, Karpov I King RD, Apple N, Fisher RE. (2002): Hyperscanning: Simultaneous fMRI during linked social interactions. *Neuroimage* 16:1159–1164.
- Montez T, Poil S-S, Jones BF, Manshanden I, Verbunt JPA, Dijk BW, van Brussaard AB, Ooyen A, van Stam CJ, Scheltens P, Linkenkaer-Hansen K (2009): Altered temporal correlations in parietal alpha and prefrontal theta oscillations in early-stage Alzheimer disease. *Proc Natl Acad Sci USA* 106:1614–1619.
- Monti MM (2011): Statistical analysis of fMRI Time-series: A critical review of the GLM approach. *Front Hum Neurosci* 5:28.
- Monto S, Palva S, Voipio J, Palva JM (2008): Very slow EEG fluctuations predict the dynamics of stimulus detection and oscillation amplitudes in humans. *J Neurosci* 28:8268–8272.
- Nikulin VV, Brismar T (2004): Long-range temporal correlations in alpha and beta oscillations: Effect of arousal level and test-retest reliability. *Clin Neurophysiol* 115:1896–1908.
- Nikulin VV, Brismar T (2005): Long-range temporal correlations in electroencephalographic oscillations: Relation to topography, frequency band, age and gender. *Neuroscience* 130:549–558.
- Nikulin VV, Hohlefeld FU, Jacobs AM, Curio G (2008): Quasi-movements: A novel motor-cognitive phenomenon. *Neuropsychologia* 46:727–742.
- Palva JM, Zhigalov A, Hirvonen J, Korhonen O, Linkenkaer-Hansen K, Palva S (2013): Neuronal long-range temporal correlations and avalanche dynamics are correlated with behavioral scaling laws. *Proc Natl Acad Sci USA* 110:3585–3590.
- Peng CK, Buldyrev SV, Havlin S (1994): Mosaic organization of DNA nucleotides. *Phys Rev E* 49:1685.
- Pessoa L, Gutierrez E, Bandettini P, Ungerleider L (2002): Neural correlates of visual working memory: fMRI amplitude predicts task performance. *Neuron* 35:975–987.
- Phillips PCB (1986): Understanding spurious regressions in econometrics. *J Econometrics* 33:311–340.
- Pritchard WS (1992): The brain in fractal time: $1/f$ -like power spectrum scaling of the human electroencephalogram. *Int J Neurosci* 66:119–129.

- Rhodes T, Turvey MT (2007): Human memory retrieval as lévy foraging. *Physica A* 385:255–260.
- Rousselet GA, Pernet CR (2012): Improving standards in brain-behavior correlation analyses. *Front Hum Neurosci* 6:119.
- Sauseng P, Klimesch W, Gerloff C Hummel F (2009): Spontaneous locally restricted EEG alpha activity determines cortical excitability in the motor cortex. *Neuropsychologia* 47:284–288.
- Smit DJ, De Geus EJ, van de Nieuwenhuijzen ME, van Beijsterveldt CE, van Baal GCM, Mansvelter HD, Linkenkaer-Hansen K (2011): Scale-free modulation of resting-state neuronal oscillations reflects prolonged brain maturation in humans. *J Neurosci* 31:13128–13136.
- Smit DJ, Linkenkaer-Hansen K, Geus EJCD (2013): Long-range temporal correlations in resting-state α oscillations predict human timing-error dynamics. *J Neurosci* 33:11212–11220.
- Theiler J, Eubank S, Longtin A (1992): Testing for nonlinearity in time series: The method of surrogate data. *Physica D* 58:77–94.
- Thiébaux HJ, Zwiers FW (1984): The interpretation and estimation of effective sample size. *J Climate Appl Meteorol* 23: 800–811.
- Van Orden GC, Holden JG, Turvey MT (2003): Self-organization of cognitive performance. *J Exp Psychol Gen* 132:331–350.
- Yule GU (1926): Why do we sometimes get nonsense-correlations between time-series? *J R Stat Soc* 89:1–63.
- Zarahn E, Aguirre GK, D’Esposito M (1997): Empirical analyses of BOLD fMRI statistics. *NeuroImage* 5:179–197.
- Zhang Y, Wang X, Bressler SL, Chen Y, Ding M (2008): Prestimulus cortical activity is correlated with speed of visuomotor processing. *J Cognit Neurosci* 20:1915–1925.