Filtering induces correlation in fMRI resting state data

Catherine E. Davey, David B. Grayden, Gary F. Egan, Leigh A. Johnston

PII: S1053-8119(12)00815-4
DOI: doi: 10.1016/j.neuroimage.2012.08.022
Reference: YNIMG 9699

To appear in: *NeuroImage*

Accepted date: 7 August 2012

Please cite this article as: Davey, Catherine E., Grayden, David B., Egan, Gary F., Johnston, Leigh A., Filtering induces correlation in fMRI resting state data, *NeuroImage* (2012), doi: 10.1016/j.neuroimage.2012.08.022

Filtering induces correlation in fMRI resting state data

Catherine E. Davey *[a,b,c], David B. Grayden[a,c], Gary F. Egan[d,e], and Leigh A. Johnston[a,b]

[a] *NeuroEngineering Laboratory, Dept. of Electrical and Electronic Engineering, University of Melbourne, Australia*

[b] *Howard Florey Institute, Florey Neuroscience Institutes, Melbourne, Australia*

[c] *NICTA Victoria Research Laboratory, Australia*

[d] *Centre for Neuroscience, University of Melbourne, Australia*

[e] *Monash Biomedical Imaging, Monash University, Australia*

May 7, 2012

Abstract

Correlation-based functional MRI connectivity methods typically impose a temporal sample independence assumption on the data. However, the conventional use of temporal filtering to address the high noise content of fMRI data may introduce sample dependence. Violation of the independence assumption has ramifications for the distribution of sample correlation which, if unaccounted for, may invalidate connectivity results. To enable the use of temporal filtering for noise suppression while maintaining the integrity of connectivity results, we derive the distribution of sample correlation between filtered timeseries as a function of the filter frequency response. Corrected distributions are also derived for statistical inference tests of sample correlation between filtered timeseries, including Fisher's z-transformation and the Student's t-test. Crucially, the proposed corrections are valid for any unknown true correlation and arbitrary filter specifications. Empirical simulations demonstrate the potential for temporal filtering to artificially induce connectivity by introducing sample dependence, and verify the utility of the proposed corrections in mitigating this effect. The importance of our corrections is exemplified in a resting state fMRI connectivity analysis: seed-voxel correlation maps generated from filtered data using uncorrected test variates yield an unfeasible number of connections to the left primary motor cortex, suggesting artificially induced connectivity, while maps acquired from filtered data using corrected test variates exhibit bilateral connectivity in the primary motor cortex, in conformance with expected results as seen in the literature.

1 Introduction

Blood-oxygenation level dependent functional MRI (BOLD-fMRI) data can be employed in fMRI connectivity analyses to gain insight into the functionality of neural networks in the brain. The BOLD signal is compromised by multiple noise sources including equipment-related noise, such as coil interference (Bandettini et al., 1998)

*Corresponding author. E-mail: cdave@ee.unimelb.edu.au. Tel: +61 3 9626 7939.

and scanner drift (Bianciardi et al., 2009), and subject-initiated noise, such as physiological interference (Biswal et al., 1995) and subject movement (Barry et al., 2010). For BOLD signals obtained at 3T, the signal of interest comprises less than 10% of the signal fluctuation, and is instead dominated by physiological noise (Kruger et al., 2001) and scanner drift (Bianciardi et al., 2009). This composition of the BOLD signal strongly motivates the use of preprocessing methods to suppress noise that, in the context of fMRI connectivity analyses, attempt to minimise artificial correlation emanating from noise processes such as synchronised cardiac and respiration signals (Birn et al., 2006). Temporal filtering is one such preprocessing method, aimed at removing noise concentrated at frequency bandwidths that do not coincide with signals of interest. Physiological noise derived from respiratory and cardiac function is concentrated at comparatively high frequencies (> 0.1Hz) (Cordes et al., 2001; Lowe et al., 1998; Thomas et al., 2002), while scanner drift is localised to frequencies below 0.01Hz (Bianciardi et al., 2009). Functional connectivity signals of interest typically subtend frequencies between 0.01-0.1Hz (Biswal et al., 1995; Cordes et al., 2001; Demirci et al., 2009; Salvador et al., 2005). Filter specification and implementation employed in fMRI connectivity literature varies considerably, the most common filter specification being 0.01-0.08Hz (Auer, 2008; Cauda et al., 2011; Uddin et al., 2009; Zhong et al., 2009). For example, Lowe et al. (1998) and Hampson et al. (2002) apply a low-pass finite impulse response (FIR) filter to remove all frequencies above 0.08Hz. Liao et al. (2010) preprocessed using a bandpass filter at 0.005-0.17Hz by forcing the appropriate frequency coefficients to zero, while Luca et al. (2006) employed a high-pass filter at 0.004Hz, implemented using a straight line fitting approach.

The use of temporal filtering as a preprocessing step modifies a voxel's frequency spectra, and hence impacts the autocorrelation profile of the signal. It is important to note that voxel timeseries may also contain intrinsic autocorrelation prior to temporal filtering (Friston et al., 1995; Purdon and Weisskoff, 1998; Lenoski et al., 2008). While Friston et al. (1995) suggested that any naturally occurring autocorrelation is negligible as a consequence of being swamped by preprocessing induced autocorrelation, Zarahn et al. (1997) demonstrated that intrinsic autocorrelation is spatially non-stationary. Consequently, autocorrelation should not be modelled using a voxel-averaged power spectra, suggesting a need for voxel-wise parametric models of autocorrelation. Lenoski et al. (2008) compared several parametric models of autocorrelation, and found that a voxel-specific autoregressive (AR) model of lag 2 accounted for the most signal autocorrelation and was sufficient to capture the range of autocorrelation found in fMRI datasets.

FMRI connectivity analyses frequently use seed-voxel correlation, in which sample correlation is calculated between a seed signal and all brain voxels (Friston, 1994; Biswal et al., 1995). The significance of a correlation estimate is typically determined via Fisher's z-transformation (Fisher, 1915) or Student's t-test (Anderson, 2003), generating a map of z-scores or t-scores, respectively. The distribution of sample correlation and, therefore, the distribution of associated significance test variates is conventionally derived using an assumption that voxel timeseries are comprised of temporally independent samples. The use of filtering in fMRI connectivity analyses violates the independence assumption by introducing autocorrelation to the timeseries and, as a consequence, statistical inference results may be invalidated. Moreover, these heterogeneous approaches to temporal filtering contain widely varying degrees of departure from the independence assumption. This has ramifications for the validity of inter-experimental comparison of results.

The variance of sample correlation proposed by Fisher (1915) is dependent on timeseries length, which dictates the degrees of freedom of the correlation estimate. By introducing autocorrelation, temporal filtering reduces the effective degrees of freedom (Kruggel et al., 2002). Friston et al. (1994) considered the problem of sample dependence in the context of fMRI activation analyses, demonstrating that the effective degrees of freedom contributing to sample correlation depends on the extent of autocorrelation in the constituent signals. Their proposed correction to z-score estimates, valid for uncorrelated signals, is calculated from the spectral densities of the constituent signals, necessitating an estimate for each signal pair. Bullmore et al. (2000) addressed the issue of sample dependence in

correlation estimates by modelling autocorrelation as a first order autoregressive process and extracting the residual of the autoregression. The ratio of the residual variance to the original variance in the timeseries was then used to weight timeseries length to obtain an estimate of the effective degrees of freedom. The efficacy of this approach depends heavily on the model of autocorrelation in the timeseries.

In this paper we derive an analytic expression for the distribution of correlation estimates acquired from temporally filtered data, from which the corrected degrees of freedom contributing to the estimate is determined. The distribution of sample correlation is expressed as a function of the complex frequency response of the filter. Crucially, the result is valid for any true correlation and arbitrary filter specifications. Furthermore, we present corrections to the significance tests of correlation, including Fisher's z-transformation and Student's t-test. We demonstrate that, in the absence of our correction, the distributions of the correlation test variates derived from filtered data do not coincide with expected distributions predicated on the independence assumption. As a consequence, intended confidence intervals are inadvertently rendered obsolete, with increased false positive rates and artificially induced correlation. The utility of our corrections for mitigating this effect and restoring confidence intervals is demonstrated using both empirical and experimental data. Our proposed corrections therefore support noise suppression via filtering, while ameliorating the problem of artificially induced connectivity. Importantly, restoring the integrity of confidence intervals supports meaningful inter-experimental comparison of results.

In this paper analytic results are discussed in the context of resting state connectivity. However, as the corrections are valid for arbitrary true correlation, they are also applicable to task-activation connectivity analyses, which is the subject of current work. The results in this paper also have implications for group level connectivity analyses given that the variance of group correlation estimates has contributions from both within-subject and between-subject variance (Penny et al., 2004).

This paper is organised as follows. In Section 2 , the necessary theoretical background is introduced. Section 3 contains our theoretical results, including the distribution of sample correlation between filtered signals, and distributions for z-score and t-score test variates derived from filtered signals. Theoretical results are verified both empirically and experimentally, the methods of which are described in Section 4, with results presented in Section 5. Section 6 presents a discussion of these results.

2 Theoretical Background

2.1 Notation

Let x_t and y_t, $t = 1, \cdots, T$, denote two bivariate normal timeseries. The covariance between x_t and y_t is defined as $\sigma_{x,y} = E\{(x - E\{x\})(y - E\{y\})\}$, and the variance of x_t, defined as the covariance of x_t with itself, is $\sigma_x^2 = E\{(x - E\{x\})(x - E\{x\})\}$. Correlation, denoted by $\rho_{x,y}$, is a measure of the instantaneous linear dependence between x_t and y_t, defined as

$$\rho_{x,y} = \frac{\sigma_{x,y}}{\sigma_x \sigma_y}. \tag{1}$$

Correlation will be used to refer to instantaneous, or lag-zero, correlation. To estimate lagged linear dependence between signals, with lag specified by τ, the cross-correlation function is used (Kumar, 2009),

$$\gamma_{x,y}(\tau) = \frac{1}{T - \tau} \sum_{t=1}^{T-\tau} x_t^* y_{t+\tau}, \tag{2}$$

where x_t^* indicates the complex conjugate of x_t.

Finally, convolution between x_t and y_t, denoted $x_t \star y_t$, is defined as

$$x_t \star y_t = \frac{1}{t}\sum_{j=1}^{t} x_j y_{t-j}. \tag{3}$$

Note that uppercase characters denote the Fourier transform of the lowercase counterparts. Values derived from filtered signals are indicated via superscript $^{(f)}$, while variates standardised to zero mean and unit variance are indicated by superscript $^{(s)}$. The spectral counterparts to standardised signal variates, not standardised themselves, will also be denoted by $^{(s)}$. Finally, to aid readability, let $\sum_k \equiv \sum_{k=0}^{T-1}$.

2.2 Statistical testing of correlation

If $\mathrm{corr}\,(x, y)$ denotes sample correlation between timeseries x_t and y_t then, for independent time points, the variance of sample correlation is given by (Hooper, 1958)

$$\sigma^2_{\mathrm{corr}(x,y)} = \frac{(1-\rho^2_{x,y})^2}{T}. \tag{4}$$

Thus, sample variance of correlation depends on the true correlation, rendering it difficult to test the significance of sample correlation values for which the true correlation is unknown. Fisher (1915) overcame this problem by proposing a transformation, known as Fisher's z-transformation:

$$\mathrm{z}\left(\mathrm{corr}\,(x, y)\right) \triangleq \frac{1}{2}\ln\frac{1+\mathrm{corr}\,(x, y)}{1-\mathrm{corr}\,(x, y)}. \tag{5}$$

The transformation to $\mathrm{z}\left(\mathrm{corr}\,(x, y)\right)$ is advantageous as the variance of $\mathrm{z}\left(\mathrm{corr}\,(x, y)\right)$ is independent of true correlation and asymptotically normal with sample size. The significance of a correlation estimate can then be tested using (Fisher, 1915)

$$\mathrm{z}\left(\mathrm{corr}\,(x, y)\right) \overset{a}{\sim} \mathcal{N}\left(\mathrm{z}\left(\rho_{x,y}\right), \frac{1}{T-1}\right), \tag{6}$$

where $\mathrm{z}\left(\rho_{x,y}\right)$ is Fisher's z-transformation of the true correlation and $\overset{a}{\sim}$ indicates that $\mathrm{z}\left(\mathrm{corr}\,(x, y)\right)$ is asymptotically normal with sample size

An alternative method employed for testing the significance of sample correlation, assuming a null hypothesis of zero correlation, is to transform $\mathrm{corr}\,(x, y)$ into a Student's t variate (Kumar, 2009):

$$\mathrm{t}\left(\mathrm{corr}\,(x, y)\right) = \mathrm{corr}\,(x, y)\sqrt{\frac{T-1}{1-\mathrm{corr}\,(x, y)^2}} \sim \boldsymbol{t}\,(T-1)\,. \tag{7}$$

3 Theoretical Results

We derive an analytic expression for the distribution of sample correlation between filtered signals, and show that the variance of sample correlation is a function of the complex frequency response of the filter. We subsequently derive corrected expressions for statistical inference tests of sample correlation, including both Fisher's z-transformation and the Student's t-test. In so doing, we demonstrate that temporal filtering induces artificial correlation by reducing the effective degrees of freedom of the estimate. Note that derivations for all equations in Section 3.1 can be found in Appendix A, while derivations for all equations in Section 3.2 are detailed in Appendix B.

3.1 Filtering induces correlation

Temporal filtering employed as a preprocessing step prior to fMRI connectivity analysis typically applies an identical filter to all voxels; consequently, the distribution of sample correlation is first derived under the assumption that an identical filter is applied to each timeseries. It is assumed that the timeseries are white, Gaussian signals containing arbitrary contemporaneous correlation.

Initially suppose that x_t and y_t are identically filtered using a filter described by the complex frequency response f_k for all frequencies k. Our first result is that the distribution of sample correlation between the filtered signals, $x_t^{(f)}$ and $y_t^{(f)}$, is

$$\operatorname{corr}\left(x^{(f)}, y^{(f)}\right) \overset{a}{\sim} \mathcal{N}\left(\rho_{x,y}, \frac{\left(1-\rho_{x,y}^2\right)^2}{\kappa}\right), \tag{8}$$

where the corrected degrees of freedom between filtered signals, κ, is

$$\kappa \triangleq \frac{\left(\sum_k f_k^2\right)^2}{\sum_k f_k^4}. \tag{9}$$

An upper bound on the corrected degrees of freedom is

$$\kappa \leq T. \tag{10}$$

The upper bound for the corrected degrees of freedom is obtained if and only if power is constant across frequencies, implying that the signal is white and, therefore, samples are independent. Eq. (10) supports the intuitive interpretation of κ as a corrected degrees of freedom. Thus, temporal filtering of timeseries may reduce the degrees of freedom contributing to a correlation estimate.

The result in Eq. (8) generalises easily to the case in which signals are filtered multiple times. The concept of multiple filters can also be employed to model sample dependence existing prior to temporal filtering. Where there are a total of n filters applied to each signal, and $f_{i,k}$ denotes the complex response of filter i at frequency k, the net complex frequency response of the filters can be determined by

$$f_k = \prod_{i=1}^{n} f_{i,k}. \tag{11}$$

The corrected degrees of freedom, Eq. (8), can then be determined as usual.

The distribution of sample correlation, Eq. (8), was derived assuming that signals were filtered using a common filter or common multiple filters. This assumption can be relaxed and the distribution of sample correlation determined for the case where signals are filtered using different filters. Let signals x_t and y_t be temporally filtered by filters specified by complex frequency responses $f_{x,k}$ and $f_{y,k}$, respectively, for all frequencies, k. Our result is that the distribution of sample correlation between the filtered signals, $x_t^{(f_x)}$ and $y_t^{(f_y)}$, is

$$\operatorname{corr}\left(x^{(f_x)}, y^{(f_y)}\right) \overset{a}{\sim} \mathcal{N}\left(\rho_{x,y}\frac{\sum_k f_{x,k}f_{y,k}}{\sqrt{\sum_k f_{x,k}^2 \sum_k f_{y,k}^2}}, \left(1-\rho_{x,y}^2\right)^2 \frac{\sum_k f_{x,k}^2 f_{y,k}^2}{\sum_k f_{x,k}^2 \sum_k f_{y,k}^2}\right). \tag{12}$$

Note that filtering the timeseries with disparate filters modifies the expected sample correlation from the original true correlation, $\rho_{x,y}$.

3.2 Correcting for induced correlation

Statistical inference testing of correlation estimates is crucial to identify significant functional connections. The distribution of a significant test variate for correlation depends on the distribution of sample correlation. The result in Eq. (8) can be utilised to correct statistical testing of correlation estimates between filtered signals. We derive corrected distributions for Fisher's z-transformation, $\mathrm{z}\left(\mathrm{corr}\left(x^{(f)}, y^{(f)}\right)\right)$ and the Student's t-test variate, $\mathrm{t}\left(\mathrm{corr}\left(x^{(f)}, y^{(f)}\right)\right)$.

Again assuming that x_t and y_t are identically filtered, using a filter described by the complex frequency response f_k, at frequency k, our result is that the distribution of Fisher's z-transformation of $\mathrm{corr}\left(x^{(f)}, y^{(f)}\right)$ is

$$\mathrm{z}\left(\mathrm{corr}\left(x^{(f)}, y^{(f)}\right)\right) \sim \mathcal{N}\left(\mathrm{z}\left(\rho_{x,y}\right), \frac{1}{\kappa}\right). \tag{13}$$

Thus, for identical filters the mean is unaffected but, as a consequence of the upper bound on κ, Eq. (10) temporal filtering of white signals necessarily increases the variance of sample correlation. The upper bound is obtained only in the case of white signals.

The derivation of this result and the following equations are in Appendix B. For an ideal filter defined by a low cut-off frequency, f_l, and a high cut-off frequency, f_h, we demonstrate that the distribution of Fisher's z-transformation of sample correlation reduces to

$$\mathrm{z}\left(\mathrm{corr}\left(x^{(f)}, y^{(f)}\right)\right) \overset{a}{\sim} \mathcal{N}\left(\mathrm{z}\left(\rho_{x,y}\right), \frac{1}{2T_s(f_h - f_l)T}\right), \tag{14}$$

where T_s denotes the sampling time for acquisition. The corrected degrees of freedom is, therefore,

$$\kappa = 2T_s(f_h - f_l)T. \tag{15}$$

Recall that for a null hypothesis of zero true correlation, sample estimates can be tested for significance using Student's t-test. We show that sample correlation between filtered signals can be transformed to a Student's t-test variate by

$$\mathrm{t}\left(\mathrm{corr}\left(x^{(f)}, y^{(f)}\right)\right) = \mathrm{corr}\left(x^{(f)}, y^{(f)}\right) \sqrt{\frac{\kappa}{\left(1 - \mathrm{corr}\left(x^{(f)}, y^{(f)}\right)^2\right)}} \tag{16}$$

with distribution

$$\mathrm{t}\left(\mathrm{corr}\left(x^{(f)}, y^{(f)}\right)\right) \sim \boldsymbol{t}\left(\frac{\left(\sum_k |f_k|^2 \star |f_k|^2\right)^2}{2\sum_k (|f_k|^2 \star |f_k|^2)^2}\right). \tag{17}$$

4 Methods

4.1 Simulated data

MATLAB was used to generate simulation datasets, each containing 2000 pairs of timeseries from a bivariate normal distribution, characterised by timeseries length $T \in [500, 10000]$, variance $\sigma^2 \in [1, 100]$, and correlation $\rho \in (-1, 1)$. Correlation induced between signals was independent of frequency and a sampling frequency of $T_s = 1$ second was used.

A set of filters, containing both realistic and unrealistic filter designs, was constructed to test the effect of different power distributions amongst signal frequencies on κ. FIR filters were included in the filter set as realistic filters

with linear phase response. AR filters were incorporated as fMRI signals are known to contain autocorrelation. Step and ramp filters were impractical filter designs that were included to examine the generality of results. The complex response vector of each filter was scaled by $s \in \{0.5, 1, 2, 4, 8\}$ to determine the impact of modifying signal power and, in the case of multiple filters being applied, the net effect of multiple filters with differing scale factors.

FIR filters were parameterised by low cut-off frequency $f_l \in [0, 0.5)$Hz, and high cut-off frequency $f_h \in (0, 0.5]$Hz, such that $f_l < f_h$, and scale of amplitude response, s. AR filters were generated using a random lag order $p \in \{0, 10\}$, and a randomised coefficient vector $\mathbf{b}$, tested for stability. Step filters were constructed by setting the frequency response between f_l and f_h to three different amplitudes, $\frac{s}{3}, \frac{s}{2}, s$, at equal divisions between f_l and f_h. The amplitude response of ramp filters increased linearly between f_l and f_h from 0 to s. An example of each filter type is shown in Fig. 1. For each filter specification, a frequency response vector was generated, $\mathbf{f} = [f_0, \cdots, f_T]'$, from which the corrected degrees of freedom, κ, was calculated.

To determine the impact of filtering on both white and non-white signals, and to verify the result for multiple filters, Eq. (11), random combinations of up to three filters were applied to datasets. Where a single filter was employed, it was applied to each timeseries after inducing correlation between signal pairs. For the case of two filters, one was applied prior to correlation being induced between timeseries pairs and the other was applied after correlation induction. Finally, in the case of three filters, one was applied prior to correlation being induced between timeseries and the other two were applied after correlation induction. The analytic expression for the corrected variance of sample correlation, Eq. (8), was evaluated using the net complex response of the multiple filters according to Eq. (11), where necessary. Sample correlation was calculated for each signal pair in the dataset and the variance across the sample correlation estimates determined. The empirical and analytic estimates for sample correlation were plotted against a sum of the filter's normalised power spectral density.

Corrected distributions for the sample correlation test statistics were verified by generating empirical distributions of the test statistics with and without the filtering correction. For each simulation dataset sample correlation between timeseries pairs was recorded both before and after filtering with random combinations of up to three filters. Fisher's z-transformation was applied and the resulting z-scores standardised to enable comparison. The z-scores generated from unfiltered data were standardised using the distribution predicated on independent samples, Eq. (6), while z-scores generated from filtered data were standardised using the uncorrected distribution and the corrected distribution that accounts for filtering, Eq. (13). A histogram was generated for each case. The same process was employed to generate empirical distributions for the t-test using the uncorrected test statistic, Eq. (7), and the test variate adjusted for filtering, Eq. (16), using the corrected distribution, Eq. (17).

4.2 Experimental data

Two healthy subjects were scanned on a 3T Siemens Tim TRIO MRI scanner using two different procedures: 1) resting state BOLD echo planar imaging (EPI), and 2) a motor task performance BOLD EPI. For both procedures, we collected 219 volumes for each participant (repetition time = 1600 ms; echo time = 20 ms; flip angle = 90°; 24 trans-axial slices, each 5.5mm thick, matrix = 64×64, field of view [FOV] = 200×200 mm^3; acquisition voxel size=$3.125 \times 3.125 \times 5.5$ mm^3). The motor task was a block design containing alternating periods of 30s each. During the active period, the subjects were instructed to press either the left or right button, according to an arrow displayed on the screen that changed randomly every second. During the baseline period, subjects were instructed to focus on a crosshair at the screen's centre point.

All images were motion corrected and smoothed using a 3D isotropic Gaussian 6mm kernel, applied using MATLAB. A conventional general linear model (GLM) activation analysis (Friston et al., 1994) was applied to the motor task data; the maximum t-statistic in the left motor cortex (LMC) region was used to identify the LMC. An average

hemodynamic timeseries was created for the LMC of each subject by averaging across voxels in the region of interest (ROI).

For each subject, seed-voxel correlation maps were calculated using the averaged LMC as the seed. A low-pass FIR filter with cut-off, $f_h = 0.1$Hz, was then applied, and the correlation maps re-calculated. For the unfiltered data, voxels significantly connected to the LMC were identified by thresholding the z-scores based on the standard distribution derived for independent samples, Eq. (6), with $\alpha < 0.01$ (Bonferroni corrected). Z-scores derived from filtered data were thresholded using both the uncorrected distribution, and the distribution corrected for low-pass filtering using Eq. (13), also with $\alpha < 0.01$ (Bonferroni corrected). Additionally, using Eq. (11), corrections were calculated for the net effect of both low-pass filtering and intrinsic autocorrelation, where the latter was estimated using a voxel-specific AR(2) process.

5 Results

The theoretical results presented in Section 3 established analytic expressions for the distribution of sample correlation between filtered timeseries based on the complex frequency responses of the filters. Analytic expressions were also provided for the test variates of sample correlation of filtered timeseries, using both Fisher's z-transformation and Student's t-test. Empirical results validating the corrections are now presented, with experimental results demonstrating the significance of the corrections in connectivity analyses.

5.1 Empirical results

The variance of sample correlation was recorded after applying each filter type to each simulation dataset. Additionally, for each simulation, the variance of sample correlation predicted by the analytic correction proposed in Eq. (8) was recorded. The simulation results (Fig. 2) demonstrate that the sample variance of correlation between filtered signals is accurately predicted by the analytic expression in Eq. (8), derived from the filter's complex amplitude response. Furthermore, the correction is clearly valid for all true correlation values. The figure also indicates that the variance of sample correlation becomes arbitrarily large as the remaining filter power becomes infinitesimally small.

Histograms of correlation test variates derived from both unfiltered and filtered Gaussian noise were generated to empirically compare the distributions of samples. Fig. 3 shows an example result generated using a true correlation of $\rho_{x,y} = 0.3$. All true correlation values generated histograms centred at 0 (data not shown) since z-scores and t-scores were standardised, and the expected value of sample correlation is unchanged when filtering signals with a common filter, Eq. (8). Standardised and corrected test variates derived from filtered timeseries have the same distribution as standardised test variates acquired from unfiltered timeseries, and conform with the standard normal distribution, indicating that the samples were standardised using the true distributions of the test variates. The uncorrected empirical distribution derived from filtered datasets show clearly the impact of temporal filtering; the measures are characterised by visibly increased variance and no longer conform to the standard normal distribution from which they were drawn, so that thresholding to identify significant sample correlation values no longer yields meaningful results. An intended z-score confidence interval of 95% is reduced to 60% if uncorrected and predicated on the violated assumption of sample independence. Conversely, thresholding the corrected measures derived from filtered data successfully restores the intended confidence interval to 95%.

5.2 Experimental results

To examine the impact of our correction in the context of fMRI connectivity, seed-voxel correlation maps were computed for two subjects using both unfiltered and filtered resting state BOLD data (Fig. 4). Thresholded maps generated from unfiltered data and significance tested using z-scores (Fig. 4A,I) and t-scores (Fig. 4E,M) show marked variation both across inference tests and between subjects. Connectivity maps generated from filtered data and significance tested using uncorrected measures show an inordinate proportion of voxels connected to the LMC (Fig. 4B,F,J,N) and exhibit substantial differences between subjects although improved consistency across significance test. Connectivity maps generated from filtered data and significance tested using our corrections for low-pass filtering contain a more realistic proportion of significant voxels connected to the LMC seed, with connectivity in closer conformance with the literature (Fig. 4C,G,K,O) (Vollmar et al., 2011; Xiong et al., 1999; Park et al., 2011), and exhibit improved consistency in connectivity both across the inference tests and between subjects. Finally, the effect of including intrinsic autocorrelation was examined for filtered data by calculating the net correction for both temporal filtering and intrinsic autocorrelation modelled as a voxel-specific, naturally occurring AR(2) filter (Fig. 4D,H,L,P). Connectivity maps containing corrections for intrinsic autocorrelation show a slightly reduced number of significant voxels when compared with those corrected only for temporal filtering, and continue to exhibit LMC connectivity in agreement with the literature.

6 Discussion

BOLD data is notoriously noisy, promoting the emergence of preprocessing techniques in fMRI connectivity analyses. Temporal filtering is a preprocessing technique that aims to reduce power at frequencies containing predominantly noise, thereby improving the signal to noise ratio (SNR). However, a secondary effect of temporal filtering, crucial in the context of fMRI connectivity, is its impact on the distribution of sample correlation. The expression for the variance of sample correlation, proposed by Fisher (1915) is founded on an assumption of sample independence and, necessarily, an absence of autocorrelation in the constituent signals. The process of filtering violates this assumption by introducing autocorrelation to the signals. We have demonstrated that temporal filtering reduces the degrees of freedom contributing to correlation estimates, thus increasing the associated variance. We have addressed this issue by deriving an analytic expression for the post-filtering distribution of sample correlation as a function of the complex frequency response of the filter.

As the distribution of sample correlation governs the distribution of its test variates, z-scores and t-scores are also impacted by temporal filtering. We have derived corrected distributions for both z-score and t-score estimates as a function of the filter frequency response, thereby restoring integrity to statistical inference results acquired from filtered datasets. If the filter-induced change in distribution is not accounted for, the consequent increase in variance will inflate the false positive rate, artificially inducing correlation.

As the extent of filtering and hence departure from sample independence increases, the expression for the variance of sample correlation proposed by Fisher (1915) becomes increasingly inaccurate. BOLD data are characterised by a narrow frequency band containing the signal of interest that renders fMRI correlation estimates vulnerable to significant deviation from the expected distribution. Furthermore, the absence of a filtering standard, in conjunction with diverse frequency spectra induced by differing sample rates, impedes comparison of results across different experimental configurations. Use of our corrections for significance testing ensures confidence intervals are correct and enables comparison of results across different experimental conditions.

The composition of BOLD signals is known to be dominated by noise, components of which are isolated in specific frequency bands. Consequently, BOLD signals may not exhibit white frequency spectra prior to temporal filtering.

There has been significant research into the estimation of intrinsic autocorrelation in BOLD fMRI data (Friston et al., 1995; Zarahn et al., 1997; Purdon and Weisskoff, 1998; Lenoski et al., 2008). Zarahn et al. (1997) demonstrated that naturally occurring autocorrelation in BOLD data is spatially non-stationary, which disallows averaging voxel power spectra to estimate intrinsic autocorrelation globally. Lenoski et al. (2008) compared different parametric models of autocorrelation, and found that a locally fit AR(2) model accounted for the most voxel autocorrelation, and therefore is the method we have applied in this study.

Intrinsic autocorrelation can be considered a naturally occurring filter, and can therefore be accounted for using Eq. (11) when significance testing sample correlation estimates. This results states that, in the case of multiple filters, the distribution of sample correlation is determined by the net filter gain. The remark is verified empirically in Section 5, which further demonstrates that the correction remains valid if a filter is applied prior to correlation being induced. The implications of this result are two-fold. First, filters can be employed to model temporal structure in the original, unprocessed data, allowing for relaxation of the independence assumption for BOLD signals. Second, the corrections are valid both when decimated frequencies contain only noise, and when filtering extracts a signal of interest.

The importance of our corrections is exemplified in application to experimental resting state data. Inconsistency of the correlation maps between the unfiltered test statistics suggests that noise is corrupting the results, advocating the need for temporal filtering to suppress noise. However, while the uncorrected test statistics acquired from filtered data show improved consistency across the test types, the improbable proportion of significant voxels is not reflective of expected results, but instead suggests severely compromised confidence intervals (Biswal et al., 1995; Xiong et al., 1999). The correlation maps derived from the corrected test statistics show improved consistency across both inference tests and subjects, and conform to established expectations (Wu et al., 2011; Vollmar et al., 2011). Correcting for both temporal filtering and intrinsic autocorrelation produces slightly fewer significant connections than correcting for temporal filtering alone. However, the change is only minor, suggesting that low-pass filtering removes much of the low frequency autocorrelation.

Friston et al. (1994) recognised the significance of autocorrelation in calculation of a corrected degrees of freedom for sample correlation. However, they estimate the variance of sample lag-0 correlation across multiple instances using the variance of cross-correlation across multiple lags. This is only valid for signals containing zero correlation at all lags. Indeed, Fisher (1915) showed that the variance of sample correlation should diminish with increasing strength of correlation, while the variance of the cross-correlation function actually increases with true correlation. This difference renders the correction proposed by Friston et al. (1994) incompatible with Fisher's z-transformation, which is applied to sample correlation to generate a test variate independent of true correlation.

The complex responses of a finite impulse response filter and an infinite impulse response filter can be approximated by an ideal filter response, described by its low and high cut-off frequencies. In this case, the corrected degrees of freedom can be determined simply from the filter cut-off frequencies, and the variance of Fisher's z-score reduces to the simple expression described in Eq. (14). This result is particularly useful in the context of fMRI connectivity, where the use of such filters is prevalent.

The impact of temporal filtering on the distribution of sample correlation estimates has implications for group analyses also, given that total variance is a function of within-subject variance (Penny et al., 2004). An in-depth examination of the impact on group analyses is beyond the scope of this paper and is the subject of current work.

7 Conclusion

In this paper we have demonstrated that the fMRI preprocessing step of temporal filtering increases the sample variance of correlation, thereby inflating the false positive rate and artificially inducing connectivity. We have derived an analytic expression for the variance of sample correlation as a function of the filter frequency response. Furthermore, we have established corrected expressions for the z-score and t-score test variates of sample correlation, restoring confidence intervals and integrity of results. We advocate employing the proposed correction to ameliorate the problem of artificially induced connectivity while retaining the improved noise suppression acquired from temporal filtering.

A Mathematical derivations for the distribution of sample correlation

Proof of Eqs. (8) and (9). A necessary first step toward deriving the sample correlation between filtered signals is to determine expressions for the spectral constituents of both the unfiltered and filtered standardised timeseries.

Timeseries x_t and y_t can be written as the Fourier transform of independent and identically distributed (iid) Gaussian frequency components,

$$\begin{aligned} x_t &= \frac{1}{T}\sum_k X_k e^{2\pi i \frac{k}{T} t}, \\ y_t &= \frac{1}{T}\sum_k Y_k e^{2\pi i \frac{k}{T} t}, \end{aligned} \tag{18}$$

where each frequency component is distributed as (Fuller, 1996)

$$\begin{aligned} X_k &\sim \mathcal{N}\left(0, T\sigma_x^2\right), \\ Y_k &\sim \mathcal{N}\left(0, T\sigma_y^2\right). \end{aligned} \tag{19}$$

The temporal variates, Eq. (18), can be transformed to standard normal variates by

$$\begin{aligned} x_t^{(s)} &\triangleq \frac{x_t}{\sigma_x} \sim \mathcal{N}(0,1) \\ y_t^{(s)} &\triangleq \frac{y_t}{\sigma_y} \sim \mathcal{N}(0,1) \end{aligned} \tag{20}$$

so that, from Eq. (19),

$$\begin{aligned} X_k^{(s)} &\triangleq \frac{X_k}{\sigma_x} \sim \mathcal{N}(0,T) \\ Y_k^{(s)} &\triangleq \frac{Y_k}{\sigma_y} \sim \mathcal{N}(0,T). \end{aligned} \tag{21}$$

If a filter, described by complex frequency response f_k, at each frequency k, is applied to both x_t and y_t, the resulting filtered signals are

$$\begin{aligned} x_t^{(f)} &\triangleq \frac{1}{T}\sum_k X_k f_k e^{2\pi i \frac{k}{T} t} \\ y_t^{(f)} &\triangleq \frac{1}{T}\sum_k Y_k f_k e^{2\pi i \frac{k}{T} t}. \end{aligned} \tag{22}$$

Note that to preserve realness of the temporal signals, the imaginary frequency components of the filter are assumed to be conjugate symmetric around 0.

Employing the distribution of the spectral variates, Eq. (19), variance of the filtered signal sample at time t, $x_t^{(f)}$, can be expressed as

$$\sigma^2_{x^{(f)}} = \frac{1}{T^2}\sum_k \left(\sigma^2_{X_k^{(f)} e^{2\pi i \frac{k}{T} t}}\right), \tag{23}$$

where $\sigma^2_{X_k^{(f)} e^{2\pi i \frac{k}{T} t}}$ denotes the variance of the spectral variate at frequency k, weighted according to the inverse Fourier transform. Letting $\mathcal{R}$ and $\mathcal{I}$ denote real and imaginary components, respectively, Eq. (23) can be rearranged

so that

$$\begin{aligned}\sigma^2_{x^{(f)}} &= \frac{1}{T^2}\sum_k \left(\sigma^2_{\mathcal{R}\{X_k^{(f)}\}} + \sigma^2_{\mathcal{I}\{X_k^{(f)}\}}\right)\\ &= \frac{1}{T^2}\sum_k \left({f_k}^* f_k \sigma^2_{\mathcal{R}\{X_k\}} + {f_k}^* f_k \sigma^2_{\mathcal{I}\{X_k\}}\right)\\ &= \frac{1}{T^2}\sum_k \left(f_k^2 \sigma^2_{\mathcal{R}\{X_k\}} + f_k^2 \sigma^2_{\mathcal{I}\{X_k\}}\right),\end{aligned} \tag{24}$$

since the conjugate symmetry of the signals requires the imaginary components to sum to zero. Furthermore, real and imaginary components of the spectral variates have equal variance (Millioz and Martin, 2010) so that

$$\begin{aligned}\sigma^2_{x^{(f)}} &= \frac{1}{T^2}\sum_k \left(\frac{f_k^2}{2}\sigma^2_{X_k} + \frac{f_k^2}{2}\sigma^2_{X_k}\right)\\ &= \frac{\sigma_x^2}{T}\sum_k f_k^2,\end{aligned} \tag{25}$$

Similarly, for $y_t^{(f)}$,

$$\sigma^2_{y^{(f)}} = \frac{\sigma_y^2}{T}\sum_k f_k^2. \tag{26}$$

Therefore, standardising the filtered timeseries results in

$$\begin{aligned}x_t^{(f,s)} &= \frac{x_t^{(f)}}{\sigma_x\sqrt{\frac{\sum_k f_k^2}{T}}}\\ y_t^{(f,s)} &= \frac{y_t^{(f)}}{\sigma_y\sqrt{\frac{\sum_k f_k^2}{T}}},\end{aligned} \tag{27}$$

which gives rise to the complementary spectral variates

$$\begin{aligned}X_k^{(f,s)} &= \frac{X_k f_k}{\sigma_x\sqrt{\frac{\sum_k f_k^2}{T}}}\\ Y_k^{(f,s)} &= \frac{Y_k f_k}{\sigma_y\sqrt{\frac{\sum_k f_k^2}{T}}}.\end{aligned} \tag{28}$$

For unfiltered bivariate normal random variables sample correlation, $\mathrm{corr}\,(x,y)$, is calculated as

$$\begin{aligned}\mathrm{corr}\,(x,y) &= \frac{\mathrm{cov}\,(x,y)}{\mathrm{std}\,(x)\,\mathrm{std}\,(y)}\\ &= \frac{\sum_{t=1}^{T}(x_t - \mathrm{M}\,(x))(y_t - \mathrm{M}\,(y))}{\sqrt{\sum_{t=1}^{T}(x_t - \mathrm{M}\,(x))^2 \sum_{t=1}^{T}(y_t - \mathrm{M}\,(y))^2}},\end{aligned} \tag{29}$$

where $\mathrm{M}\,(x)$ and $\mathrm{M}\,(y)$ denote the sample means of x_t and y_t, respectively. Sample correlation can be equivalently expressed using standardised variates by a simple rearrangement of Eq. (29) (Whittaker, 2009),

$$\begin{aligned}\mathrm{corr}\,(x,y) &= \frac{1}{T}\sum_{t=1}^{T}\frac{(x_t - \mathrm{M}\,(x))}{\mathrm{std}\,(x)}\frac{(y_t - \mathrm{M}\,(y))}{\mathrm{std}\,(y)}\\ &= \frac{1}{T}\sum_{t=1}^{T} x_t^{(s)} y_t^{(s)},\end{aligned} \tag{30}$$

where $x_t^{(s)}$ and $y_t^{(s)}$ denote the standard normal counterparts of x_t and y_t, respectively, such that $x_t^{(s)} = \frac{(x_t - \mathrm{M}(x))}{\mathrm{std}(x)}$, and $y_t^{(s)} = \frac{(y_t - \mathrm{M}(y))}{\mathrm{std}(y)}$. Correlation is equivalent to cross-correlation, Eq. (2), between standardised variates at lag zero,

$$\rho_{x,y} = \gamma_{x^{(s)}y^{(s)}}(0). \tag{31}$$

Employing the convolution theorem, the Fourier transform of cross-correlation satisfies (Kumar, 2009)

$$\Gamma_{x,y}(k) = \frac{1}{T} X_k^* Y_k, \qquad k = 1, \cdots, T, \tag{32}$$

where X_k and Y_k denote the values for frequency k, of the Fourier transforms of the vectors $\mathbf{x} = [x_1, \cdots, x_T]'$ and $\mathbf{y} = [y_1, \cdots, y_T]'$, respectively. Consequently, application of the inverse Fourier transform requires that

$$\gamma_{x,y}(\tau) = \frac{1}{T^2} \sum_{k=0}^{T-1} X_k^* Y_k e^{2\pi i \frac{k}{T}\tau}. \tag{33}$$

To derive the distribution of sample correlation between filtered variates express sample correlation as lag-0 cross-correlation (Eqs. (31) and (33)) between Eq. (28), which gives

$$\begin{aligned}
\mathrm{corr}\left(x^{(f)}, y^{(f)}\right) &= \frac{1}{T^2} \sum_k {X_k^{(f,s)}}^* Y_k^{(f,s)} e^{2\pi i \frac{k}{T} 0} \\
&= \frac{1}{T^2} \sum_k \left(\frac{\left(X_k f_k\right)^*}{\sigma_x \sqrt{\frac{\sum\limits_k f_k^2}{T}}} \frac{Y_k f_k}{\sigma_y \sqrt{\frac{\sum\limits_k f_k^2}{T}}} \right) \\
&= \frac{1}{T \sum\limits_k f_k^2} \sum_k \left(\frac{X_k^*}{\sigma_x} \frac{Y_k}{\sigma_y} {f_k}^* f_k \right) \\
&= \frac{1}{T \sum\limits_k f_k^2} \sum_k \left(\frac{X_k^*}{\sigma_x} \frac{Y_k}{\sigma_y} f_k^2 \right),
\end{aligned} \tag{34}$$

where the last line follows from the conjugate symmetry of the filtered signals.

Although the product terms, $X_k^* Y_k$, Eq. (34), have a Laplacian distribution (see Appendix C), the central limit theorem dictates that the sum tends to a Gaussian distribution. To describe the distribution of sample correlation, we therefore only require the mean and variance as sufficient statistics of the Gaussian distribution.

Application of Eq. (21) to Eq. (34), in conjunction with the iid attribute of the unfiltered spectral components, yields the expected value of $\mathrm{corr}\left(x^{(f)}, y^{(f)}\right)$:

$$\begin{aligned}
\mathrm{E}\left\{\mathrm{corr}\left(x^{(f)}, y^{(f)}\right)\right\} &= \frac{1}{T \sum\limits_k f_k^2} \mathrm{E}\left\{X_k^{(s)} Y_k^{(s)}\right\} \sum_k T f_k^2 \\
&= \frac{\rho_{x,y}}{T \sum\limits_k f_k^2} \sum_k T f_k^2 \\
&= \frac{\rho_{x,y}}{\sum\limits_k f_k^2} \sum_k f_k^2 \\
&= \rho_{x,y}.
\end{aligned} \tag{35}$$

To determine the variance of sample correlation, rewriting Eq. (34) using the standardised spectral components of

the serially independent signals, Eq. (21), and employing the independence of spectral variates obtains

$$\sigma^2_{\mathrm{corr}\left(x^{(f)},y^{(f)}\right)} = \frac{1}{T^2\left(\sum_k f_k^2\right)^2}\sum_k \sigma^2_{X_k^{(s)*}Y_k^{(s)}} f_k^4. \tag{36}$$

It remains to determine $\sigma^2_{X_k^{(s)}Y_k^{(s)}}$. Considering the standardised form of sample correlation, Eq. (30), we can write,

$$\sigma^2_{\mathrm{corr}(x,y)} = \frac{1}{T^2}\sum_{t=1}^{T}\sigma^2_{x^{(s)}y^{(s)}}, \tag{37}$$

since the unfiltered signals are comprised of independent temporal samples. Rearranging Eq. (37), and employing the known variance of sample correlation between unfiltered signals, Eq. (4), gives

$$\sigma^2_{x^{(s)}y^{(s)}} = \left(1-\rho^2_{x,y}\right)^2, \tag{38}$$

which implies that

$$\sigma^2_{X_k^{(s)}Y_k^{(s)}} = T^2\left(1-\rho^2_{x,y}\right)^2. \tag{39}$$

Substitution of Eq. (39) in Eq. (36) yields the sample variance of correlation between filtered signals,

$$\begin{aligned}\sigma^2_{\mathrm{corr}\left(x^{(f)},y^{(f)}\right)} &= \frac{1}{T^2\left(\sum_k f_k^2\right)^2}\sum_k T^2\left(1-\rho^2_{x,y}\right)^2 f_k^4 \\ &= \frac{\sum_k f_k^4}{\left(\sum_k f_k^2\right)^2}\left(1-\rho^2_{x,y}\right)^2,\end{aligned} \tag{40}$$

and the proof is complete.

□

Proof of Eq. (10). Recall that the corrected degrees of freedom, Eq. (9), is given by

$$\kappa \triangleq \frac{\left(\sum_k f_k^2\right)^2}{\sum_k \left(f_k^2\right)^2}.$$

Denote the power at frequency k by $p_k = f_k^2$, so that

$$\kappa = \frac{\left(\sum_k p_k\right)^2}{\sum_k p_k^2}. \tag{41}$$

Application of Jensen's inequality with identically unitary weights for all frequencies requires that (Jensen, 1906)

$$\left(\frac{\sum_k p_k}{T}\right)^2 \leq \frac{\sum_k p_k^2}{T}. \tag{42}$$

Noting that p_k is necessarily positive, rearranging gives

$$\frac{\left(\sum_k p_k\right)^2}{\sum_k p_k^2} \leq T. \tag{43}$$

Substituting f_k^2 for p_k completes the proof.

□

Proof of Eq. (11). The proof follows directly from the definition of a temporally filtered signal, Eq. (22).

□

Proof of Eq. (12). Note that this derivation follows the same procedure as the derivation of Eq. (8) and is included for completeness.

Temporal filtering of white, Gaussian timeseries x_t and y_t by disparate filters respectively defined by complex frequency responses, $f_{x,k}$ and $f_{y,k}$, for all frequencies k, generates filtered signals $x_t^{(f_x)}$ and $y_t^{(f_y)}$, respectively. The filtered signals can be expressed as

$$\begin{aligned} x_t^{(f_x)} &= \frac{1}{T}\sum_k X_k f_{x,k} e^{2\pi i \frac{k}{T} t}, \\ y_t^{(f_y)} &= \frac{1}{T}\sum_k Y_k f_{y,k} e^{2\pi i \frac{k}{T} t}. \end{aligned} \tag{44}$$

Variance of the filtered signals can be determined by application of Eq. (25) to be

$$\begin{aligned} \sigma^2_{x^{(f_x)}} &= \frac{\sigma_x^2}{T}\sum_k f_{x,k}^2 \\ \sigma^2_{y^{(f_y)}} &= \frac{\sigma_y^2}{T}\sum_k f_{y,k}^2. \end{aligned} \tag{45}$$

Therefore, standardising the filtered time series, $x_t^{(f_x)}$ and $y_t^{(f_y)}$, requires

$$x_t^{(f_x s)} = \frac{x_t^{(f_x)}}{\sigma_x \sqrt{\frac{\sum_k f_{x,k}^2}{T}}} \tag{46}$$

$$y_t^{(f_y s)} = \frac{y_t^{(f_y)}}{\sigma_y \sqrt{\frac{\sum_k f_{y,k}^2}{T}}}, \tag{47}$$

which generates the complementary spectral variates

$$\begin{aligned} X_k^{(f_x s)} &= \frac{X_k f_{x,k}}{\sigma_x \sqrt{\frac{\sum_k f_{x,k}^2}{T}}} \\ Y_k^{(f_y s)} &= \frac{Y_k f_{y,k}}{\sigma_y \sqrt{\frac{\sum_k f_{y,k}^2}{T}}}. \end{aligned} \tag{48}$$

Expressing sample correlation as a special case of cross-correlation, Eq. (31), and deriving cross-correlation from its spectral components, Eq. (33), results in

$$\begin{aligned}
\operatorname{corr}\left(x^{(f_x)}, y^{(f_y)}\right) &= \frac{1}{T^2}\sum_{k=0}^{T-1} X_k^{(f_x s)^*} Y_k^{(f_y s)} \\
&= \frac{1}{T^2}\sum_k \left(\frac{{X_k}^* {f_{x,k}}^*}{\sigma_x\sqrt{\frac{\sum_k f_{x,k}^2}{T}}} \frac{Y_k f_{y,k}}{\sigma_y\sqrt{\frac{\sum_k f_{y,k}^2}{T}}} \right) \\
&= \frac{1}{T\sqrt{\sum_k f_{x,k}^2 \sum_k f_{y,k}^2}} \sum_k \left(\frac{X_k^*}{\sigma_x}\frac{Y_k}{\sigma_y} {f_{x,k}}^* f_{y,k} \right) \\
&= \frac{1}{T\sqrt{\sum_k f_{x,k}^2 \sum_k f_{y,k}^2}} \sum_k \left(\frac{X_k}{\sigma_x}\frac{Y_k}{\sigma_y} f_{x,k} f_{y,k} \right),
\end{aligned} \tag{49}$$

where the last line results because the filtered signals are assumed to be real.

Application of the central limit theorem requires that the sum tends to a Gaussian distribution. Consequently, mean and variance are sufficient to describe the distribution of sample correlation between timeseries filtered with disparate filters.

The expected value of $\operatorname{corr}\left(x^{(f_x)}, y^{(f_y)}\right)$ can be derived from Eq. (49) by observing that the standardised spectral components of the unfiltered timeseries, Eq. (21), are iid;

$$\begin{aligned}
\mathrm{E}\left\{\operatorname{corr}\left(x^{(f_x)}, y^{(f_y)}\right)\right\} &= \frac{1}{T\sqrt{\sum_k f_{x,k}^2 \sum_k f_{y,k}^2}} \mathrm{E}\left\{X_k^{(s)} Y_k^{(s)}\right\} \sum_k \left(f_{x,k} f_{y,k}\right) \\
&= \frac{\rho_{x,y} T}{T\sqrt{\sum_k f_{x,k}^2 \sum_k f_{y,k}^2}} \sum_k \left(f_{x,k} f_{y,k}\right) \\
&= \rho_{x,y} \frac{\sum_k f_{x,k} f_{y,k}}{\sqrt{\sum_k f_{x,k}^2 \sum_k f_{y,k}^2}}.
\end{aligned} \tag{50}$$

To determine the variance of sample correlation between disparately filtered signals rewrite Eq. (49) using the standardised spectral components of unfiltered signals $X_k^{(s)}$ and $Y_k^{(s)}$, Eq. (21), as

$$\sigma^2_{\operatorname{corr}\left(x^{(f_x)}, y^{(f)}\right)} = \frac{1}{T^2\sqrt{\sum_k f_{x,k}^2 \sum_k f_{y,k}^2}} \sum_k \sigma^2_{X_k^{(s)*} Y_k^{(s)}} f_{x,k}^2 f_{y,k}^2. \tag{51}$$

Application of Eq. (39) then gives

$$\begin{aligned}
\sigma^2_{\operatorname{corr}\left(x^{(f_x)}, y^{(f_y)}[_y]\right)} &= \frac{1}{T^2 \sum_k f_{x,k}^2 \sum_k f_{y,k}^2} \sum_k T^2 \left(1-\rho_{x,y}^2\right)^2 f_{x,k}^2 f_{y,k}^2 \\
&= \left(1-\rho_{x,y}^2\right)^2 \frac{\sum_k f_{x,k}^2 f_{y,k}^2}{\sum_k f_{x,k}^2 \sum_k f_{y,k}^2},
\end{aligned} \tag{52}$$

and the proof is complete.

□

B Mathematical derivations for statistical inference tests

Proof of Eq. (13). The variance of $\mathrm{z}\left(\mathrm{corr}\left(x^{(f)}, y^{(f)}\right)\right)$ can be derived via the delta method (Oehlert, 1992), which states that if $\mathrm{corr} \sim \mathcal{N}\left(\rho, \sigma^2_{\mathrm{corr}}\right)$ then for some function $g(\mathrm{corr})$ with derivative defined at ρ, a first order approximation of its distribution is given by

$$g(\mathrm{corr}) \xrightarrow{D} \mathcal{N}\left(g(\rho), \sigma^2_{\mathrm{corr}} g'(\rho)^2\right), \tag{53}$$

where $\xrightarrow{D}$ denotes convergence in distribution.

Let

$$g(\mathrm{corr}) \triangleq \mathrm{z}\left(\mathrm{corr}\right) = \frac{1}{2}\ln\frac{1+\mathrm{corr}}{1-\mathrm{corr}} \tag{54}$$

so that

$$g'(\mathrm{corr}) = \frac{1}{1-\mathrm{corr}^2}. \tag{55}$$

Applying Eq. (8) to Eq. (53) gives

$$\begin{aligned}
\sigma^2_{\mathrm{z}\left(\mathrm{corr}\left(x^{(f)}, y^{(f)}\right)\right)} &= \sigma^2_{\mathrm{corr}\left(x^{(f)}, y^{(f)}\right)} \left(g'(\rho_{x^{(f)}, y^{(f)}})\right)^2 \\
&= \frac{\left(1-\rho^2_{x^{(f)}, y^{(f)}}\right)^2}{\kappa} \frac{1}{\left(1-\rho_{x^{(f)}, y^{(f)}}\right)^2} \\
&= \frac{1}{\kappa},
\end{aligned} \tag{56}$$

proving the variance component. From Eq. (8), the mean of $\mathrm{corr}\left(x^{(f)}, y^{(f)}\right)$ is simply $\rho_{x,y}$; transforming the mean using $g(\rho_{x,y})$, the result follows.

□

Proof of Eqs. (14) and (15). Define an ideal filter by a low cut-off frequency, f_l, and a high cut-off frequency, f_h. For a timeseries of length T denote the interval between discrete frequencies by (Atlas, 2009, p.54)

$$\Delta f = \frac{1}{T}, \tag{57}$$

so that $f_l = \Delta f k_l$ and $f_h = \Delta f k_h$, where k_l and k_h denote indices for the low and high cut-off frequencies, respectively.

Since the filter is ideal

$$f_k = \begin{cases} 0, & \text{for } k < f_l, f_h < k \\ 1, & \text{for } f_l \leq k \leq f_h. \end{cases} \tag{58}$$

so that the expression for corrected degrees of freedom between filtered signals, Eq. (9), reduces to

$$\begin{aligned}
\kappa &= \frac{\left(2\left(k_h - k_l\right)\right)^2}{2\left(k_h - k_l\right)} \\
&= 2\left(k_h - k_l\right) \\
&= \frac{2\left(f_h - f_l\right)}{\Delta f}.
\end{aligned} \tag{59}$$

The sampling interval can be expressed as (Westbrook et al., 2005, pp.77-81)

$$T_s = \frac{1}{T\Delta f}. \tag{60}$$

Substituting this into Eq. (59) gives

$$\kappa = 2T_s T \left(f_h - f_l\right), \tag{61}$$

and the proof is complete.

□

Proof of Eqs. (16) and (17). As the t-test is only valid when testing the null case of zero correlation, it suffices for this proof to consider only the case in which $x_t^{(f)}$ and $y_t^{(f)}$ are independent.

To aid readability in deriving the transformation of sample correlation to a Student's t statistic, define

$$w_t = x_t^{(f,s)} y_t^{(f,s)} \tag{62}$$

with sample mean

$$\begin{aligned} \mathrm{M}\left(w\right) &= \frac{1}{T}\sum_{t=1}^{T} w_t \\ &= \frac{1}{T}\sum_{t=1}^{T} x_t^{(f,s)} y_t^{(f,s)}, \end{aligned} \tag{63}$$

which is equivalent to sample correlation between filtered signals, $\mathrm{corr}\left(x^{(f)}, y^{(f)}\right)$, Eq. (30).

The Student's t-test statistic is acquired from the ratio of a standard normal variate to an independent χ variate. The t-test for sample correlation obtains these variables from descriptive statistics of w_t.

The standard normal variate is derived by standardising the sample mean, $\mathrm{M}\left(w\right)$. From Eq. (8), μ_w (i.e. $\mathrm{corr}\left(x^{(f)}, y^{(f)}\right)$) has an asymptotically normal distribution with mean zero, and can be standardised by,

$$\begin{aligned} \mathrm{M}\left(w\right)^{(s)} &= \frac{\mathrm{M}\left(w\right)}{\sigma_{\mu_w}} \\ &= \mathrm{M}\left(w\right)\sqrt{\kappa}. \end{aligned} \tag{64}$$

To construct the χ variate, consider the sample variance of w_t,

$$\begin{aligned} \mathrm{var}\left(w\right) &= \frac{1}{T-1}\sum_{t=1}^{T}\left(w_t^2 - \mathrm{M}\left(w\right)^2\right) \\ &= \frac{T}{T-1}\left(1 - \mathrm{M}\left(w\right)^2\right), \end{aligned} \tag{65}$$

where the last line is a consequence of Eq. (62), which describes w_t as the product of independent, standard normal variates, so that $E\{w\} = 0$ and $\sigma_w^2 = 1$. Application of Parseval's theorem (Kammler, 2000) under the null hypothesis requires that

$$\begin{aligned} \sum_{t=1}^{T} w_t^2 &= \frac{1}{T}\sum_{k} |W_k|^2 \\ \mathrm{var}\left(w\right) &= \frac{1}{T}\mathrm{var}\left(W\right), \end{aligned} \tag{66}$$

which permits the use of the sample variance of W_k to determine the signal variance for w_t. Since w_t is derived from the product of temporal variables, $x_t^{(f,s)}$ and $y_t^{(f,s)}$, Eq. (62), the convolution theorem dictates that W_k is determined by the convolution of the frequency domain counterparts, $X_k^{(f,s)}$ and $Y_k^{(f,s)}$, Eq. (28), so that

$$\begin{aligned} W_k &= X_k^{(f,s)} \star Y_k^{(f,s)} \\ &= \frac{T}{T\sigma_x\sigma_y \sum\limits_k f_k^2} \left(f_k X_k \star f_k Y_k\right) \\ &= \frac{1}{\sum\limits_k f_k^2} \sum_{j=1}^{N} f_j f_{k-j} X_k^{(s)} Y_{k-j}^{(s)}. \end{aligned} \tag{67}$$

Although each term in the sum has a Laplacian distribution (see Appendix C), application of the central limit theorem results in an asymptotic normal distribution for W_k, with zero mean and variance,

$$\begin{aligned} \mathrm{var}\left(W\right) &= \frac{1}{\sum\limits_k f_k^2} \sum_{j=1}^{N} |f_j|^2 |f_{k-j}|^2 \sigma^2_{X_j^{(s)}} \sigma^2_{Y_{k-j}^{(s)}} \\ &= \frac{T^2}{\left(\sum\limits_k f_k^2\right)^2} |f_k|^2 \star |f_k|^2. \end{aligned} \tag{68}$$

Consequently, the distribution of the variance of W_k and hence w_t (Eq. (66)) can be considered a weighted sum of χ^2 variates, which can be approximated by a χ^2 distribution with effective degrees of freedom determined by the Welch-Satterthwaite's equation (Satterthwaite, 1946; Welch, 1947),

$$\begin{aligned} d &= \frac{\left(\sum\limits_k |W_k|^2\right)^2}{\sum\limits_k |W_k|^4} \\ &= \frac{\left(\frac{T^2}{\left(\sum_k f_k^2\right)^2} \sum\limits_k |f_k|^2 \star |f_k|^2\right)^2}{2\sum\limits_k \frac{T^4}{\left(\sum_k f_k^2\right)^4} \left(|f_k|^2 \star |f_k|^2\right)^2} \\ &= \frac{\left(\sum\limits_k |f_k|^2 \star |f_k|^2\right)^2}{2\sum\limits_k \left(|f_k|^2 \star |f_k|^2\right)^2}, \end{aligned} \tag{69}$$

where the factor of $\frac{1}{2}$ is a consequence of d being an even function.

The standard normal and χ variates from which the Student's t-test statistic is constructed are required to be independent. Given that the sample mean and sample variance of a Gaussian variable are independent[1] (Whittaker, 2009), a ratio of the sample mean, $\mathrm{M}\left(w\right)$, and sample variance, $\mathrm{var}\left(w\right)$, both scaled by σ_w, satisfies the independence requirement. To this end define

$$\begin{aligned} v &\triangleq (T-1)\frac{\mathrm{var}\left(w\right)}{\sigma_w^2} \\ &= T\left(1 - \mathrm{M}\left(w\right)^2\right), \end{aligned} \tag{70}$$

since $\sigma_w^2 = 1$.

[1]The sample mean and sample variance of a Gaussian, $z_t \sim \mathcal{N}\left(\mu_z, \sigma_z^2\right)$, are independent since $E\{\mu_z(z_t - \mu_z)\} = 0$.

The Student's t variate can now be constructed from the ratio of Eqs. (64) and (70),

$$\frac{\mathrm{M}\left(w\right)^{(s)}}{\sqrt{v}} \sim \boldsymbol{t}\left(d\right). \tag{71}$$

Substituting $\mathrm{corr}\left(x^{(f)}, y^{(f)}\right)$ for $\mathrm{M}\left(w\right)$, and employing the expression derived for the degrees of freedom Eq. (69) results in

$$\begin{aligned} t\left(\mathrm{corr}\left(x^{(f)}, y^{(f)}\right)\right) &= \mathrm{corr}\left(x^{(f)}, y^{(f)}\right)\sqrt{\frac{\kappa}{1-\mathrm{corr}\left(x^{(f)}, y^{(f)}\right)^2}} \\ t\left(\mathrm{corr}\left(x^{(f)}, y^{(f)}\right)\right) &\sim t\left(\frac{\left(\sum\limits_k |f_k|^2 \star |f_k|^2\right)^2}{2\sum\limits_k \left(|f_k|^2 \star |f_k|^2\right)^2}\right). \end{aligned} \tag{72}$$

□

C Product of complex Gaussian variates

The product of two Gaussian random variables results in a product-normal variate. Assuming we take the product of $x \sim p_{\mathcal{N}}\left(x\,;0,\sigma_x^2\right)$ and $y \sim p_{\mathcal{N}}\left(y\,;0,\sigma_y^2\right)$, the resulting product normal variate, $xy \sim p_{\mathcal{PN}}\left(xy\,;\sigma_x\sigma_y\right)$, has the characteristic function $\frac{1}{(1+\sigma_x^2\sigma_y^2)^{\frac{1}{2}}}$ (Pearson et al., 1929). The sum of two identical product normal variates has the characteristic function

$$\psi = \frac{1}{(1+\sigma_x^2\sigma_y^2t^2)^{\frac{1}{2}}}\frac{1}{(1+\sigma_x^2\sigma_y^2t^2)^{\frac{1}{2}}} = \frac{1}{(1+\sigma_x^2\sigma_y^2t^2)} \sim p_{\mathcal{L}apl}\left(xy\,;0,\sigma_x\sigma_y\right). \tag{73}$$

Thus, the sum of two identical product normal variables generates a Laplacian variable.

References

Anderson, T., 2003. An introduction to multivariate statistical analysis, 3rd Edition. Wiley.

Atlas, S., 2009. Magnetic resonance imaging of the brain and spine, 4th Edition. Lippincott Williams and Wilkins.

Auer, D., 2008. Spontaneous low-frequency blood oxygenation level-dependent fluctuations and functional connectivity analysis of the "resting" brain. Magnetic Resonance Imaging 26, 1055–1064.

Bandettini, P., Jesmanowicz, A., Kylen, J., Birn, R., Hyde, J., 1998. Functional MRI of brain activation induced by scanner acoustic noise. Magnetic Resonance in Medicine 39, 410–416.

Barry, R., Williams, J., Klassen, L., Gallivan, J., Culham, J., Menon, R., 2010. Evaluation of preprocessing steps to compensate for magnetic field distortions due to body movements in bold fMRI. Magnetic Resonance Imaging 28 (2), 235–244.

Bianciardi, M., Fukunaga, M., van Gelderen, P., Horovitz, S., de Zwart, J., Shmueli, K., Duyn, J., 2009. Sources of functional magnetic resonance imaging signal fluctuations in the human brain at rest: A 7 T study. Magnetic Resonance Imaging 27 (8), 1019–1029.

Birn, R., Diamond, J., Smith, M., Bandettini, P., 2006. Separating respiratory-variation-related fluctuations from neuronal-activity-related fluctuations in fMRI. NeuroImage 31, 1536–1548.

Biswal, B., Yetkin, F., Haughton, V., Hyde, J., 1995. Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. Magnetic Resonance in Medicine 34, 537–541.

Bullmore, E., Horwitz, B., Honey, G., Brammer, M., Williams, S., Sharma, T., 2000. How good is good enough in path analysis of fMRI data? NeuroImage 11, 289–301.

Cauda, F., D'Agata, F., Sacco, K., Duca, S., Geminiani, G., Vercelli, A., 2011. Functional connectivity of the insula in the brain. NeuroImage 55, 8–23.

Cordes, D., Haughton, V., Arfanakis, K., Carew, J., Turski, P., Moritz, C., Quigley, M., Meyerand, M., 2001. Frequencies contributing to functional connectivity in the cerebral cortex in 'resting-state' data. American Journal of Neuroradiology, 1326–1333.

Demirci, O., Stevens, M., Andreasen, N., Michael, A., Liu, J., White, T., Pearlson, G., Clark, V., Calhoun, V., 2009. Investigation of relationships between fMRI brain networks in the spectral domain using ICA and Granger causality reveals distinct differences between schizophrenia patients and healthy controls. NeuroImage 46, 419–431.

Fisher, R., 1915. Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. Biometrika 10, 507–521.

Friston, K., 1994. Functional and effective connectivity in neuroimaging: A synthesis. Human Brain Mapping 2, 56–78.

Friston, K., Holmes, A., Poline, J., Grasby, P., Williams, S., Frackowiak, R., Turner, R., 1995. Analysis of fmri time series revisited. NeuroImage 2, 45–53.

Friston, K., Jezzard, P., Turner, R., 1994. Analysis of functional MRI time-series. Human Brain Mapping 1, 153–171.

Fuller, W., 1996. Introduction to statistical time series, 2nd Edition. Wiley.

Hampson, M., Peterson, B., Skudlarski, P., Gatenby, J., Gore, J., 2002. Detection of functional connectivity using temporal correlations in MR images. Human Brain Mapping 15, 247–262.

Hooper, J., 1958. The sampling variance of correlation coefficients under assumptions of fixed and mixed variates. Biometrika Trust 45 (3), 471–477.

Jensen, J., 1906. Sur les fonctions convexes et les ingalits entre les valeurs moyennes. Acta Mathematica 30 (1), 175–193.

Kammler, D., 2000. A First Course in Fourier Analysis, 22nd Edition. Prentice-Hall.

Kruger, G., Kastrup, A., Glover, G., 2001. Neuroimaging at 1.5 T and 3.0 T: Comparison of oxygenation-sensitive magnetic resonance imaging. Magnetic Resonance in Medicine 45, 595–604.

Kruggel, F., Pelegrini-Issac, M., Benali, H., 2002. Estimating the effective degrees of freedom in univariate multiple regression analysis. Medical Image Analysis 6, 63–75.

Kumar, R., 2009. Signals and systems. PHI Learning Pvt. Ltd.

Lenoski, B., Baxter, L., Karam, L., Maisog, J., Debbins, J., 2008. On the performance of autocorrelation estimation algorithms for fMRI analysis. IEEE Journal of Selected Topics in Signal Processing 2 (6), 828–838.

Liao, W., Mantini, D., Zhang, Z., Pan, Z., Ding, J., Gong, Q., Yang, Y., Chen, H., 2010. Evaluating the effective connectivity of resting state networks using conditional Granger causality. Biological Cybernetics 102, 57–69.

Lowe, M., Mock, B. J., Sorenson, J., 1998. Functional connectivity in single and multislice echoplanar imaging using resting state fluctuations. NeuroImage 7, 119–132.

Luca, M., Beckmann, C., Stefano, N., Matthews, P., Smith, S., 2006. FMRI resting state networks define distinct modes of long-distance interactions in the human brain. NeuroImage 29, 1359–1367.

Millioz, F., Martin, N., 2010. Estimation of a white gaussian noise in the short time Fourier transform based on the spectral kurtosis of the minimal statistics: application to underwater noise. In: IEEE International Conference on Acoustics, Speech, and Signal Processing. Dallas, Texas, United States, pp. 5638–5641.

Oehlert, G., 1992. A note on the delta method. The American Statistician 46 (1), 27–29.

Park, C., Chang, W., Ohn, S., Kim, S., Bang, O., Pascual-Leone, A., Kim, Y., 2011. Longitudinal changes of resting-state functional connectivity during motor recovery after stroke. Stroke 42 (5), 1357–62.

Pearson, K., Jeffery, G., Elderton, E., 1929. On the distribution of the first product moment-coefficient, in samples drawn from an indefinitely large normal population. Biometrika 21, 164–201.

Penny, W. D., Holmes, A. J., Friston, K., 2004. Random effects analysis, 2nd Edition. Academic Press.

Purdon, P., Weisskoff, R., 1998. Effect of temporal autocorrelation due to physiological noise and stimulus paradigm on voxel-level false-positive rates in fMRI. Human Brain Mapping 6, 239–249.

Salvador, R., Suckling, J., Coleman, M., Pickard, J., Menon, D., Bullmore, E., 2005. Neurophysiological architecture of functional magnetic resonance images of human brain. Cerebral Cortex 15 (9), 1332–1342.

Satterthwaite, F., 1946. An approximate distribution of estimates of variance components. Biometrics Bulletin 2, 110–114.

Thomas, C., Harshman, R., Menon, R., 2002. Noise reduction in BOLD-based fMRI using component analysis. NeuroImage 17, 1521–1537.

Uddin, L., Kelly, A., Biswal, B., Castellanos, F., Milham, M., 2009. Functional connectivity of default mode network components: Correlation, anticorrelation, and causality. Human Brain Mapping 30, 625–637.

Vollmar, C., OMuircheartaigh, J., Barker, G., Symms, M., Thompson, P., Kumari, V., Duncan, J., Janz, D., Richardson, M., Koepp, M., 2011. Motor system hyperconnectivity in juvenile myoclonic epilepsy: a cognitive functional magnetic resonance imaging study. Brain 134 (6), 1710–1719.

Welch, B., 1947. The generalization of "Student's" problem when several different population variances are involved. Biometrika 34, 28–35.

Westbrook, C., Roth, C., Talbot, J., 2005. MRI in practice, 3rd Edition. Blackwell Publishing.

Whittaker, J., 2009. Graphical models in applied multivariate statistics, 3rd Edition. Wiley Publishing.

Wu, T., Long, X., Wang, L., Hallett, M., Zang, Y., Li, K., Chan, P., 2011. Functional connectivity of cortical motor areas in the resting state in parkinson's diseas. Human Brain Mapping 32 (9), 1443–1457.

Xiong, J., Parsons, L., Gao, J., Fox, P., 1999. Interregional connectivity to primary motor cortex revealed using MRI resting state images. Human Brain Mapping 8, 151–156.

Zarahn, E., Aguirre, G. K., DEsposito, M., 1997. Empirical analyses of bold fmri statistics. NeuroImage 5 (3), 179–197.

Zhong, Y., Wang, H., Lu, G., Zhang, Z., Jiao, Q., Liu, Y., 2009. Detecting functional connectivity in fMRI using PCA and regression analysis. Brain Topography 22, 134–44.

List of Figures

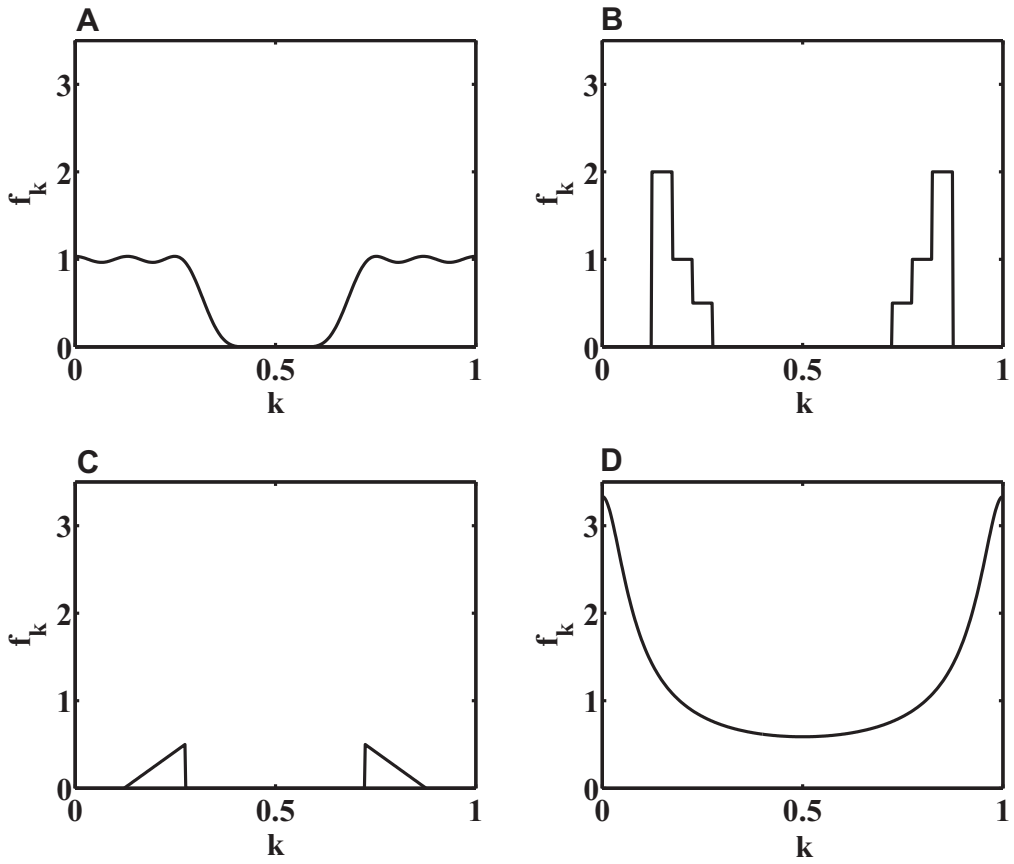

Figure 1: Amplitude response of example filters from each filter class. **A:** FIR filter with cut-off frequencies $f_l = 0$, $f_h = 0.25$ and scaling parameter $s = 1$. **B:** Step filter with cut-off frequencies $f_l = 0.125$, $f_h = 0.275$ and scale parameter $s = 2$. **C:** Ramp filter with cut-off frequencies $f_l = 0.125$, $f_h = 0.275$ and scale parameter $s = 0.5$. **D:** AR(1) filter with coefficient of 0.9.

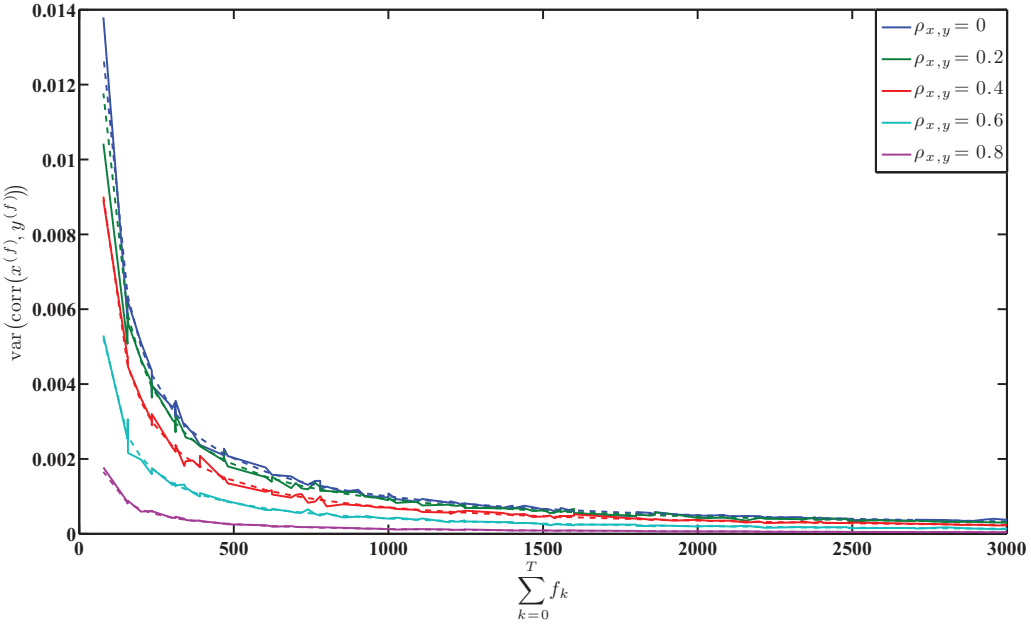

Figure 2: Variance of sample correlation estimated from simulations (solid lines), and predicted using the filter's complex response (dashed lines), for various true correlation values.

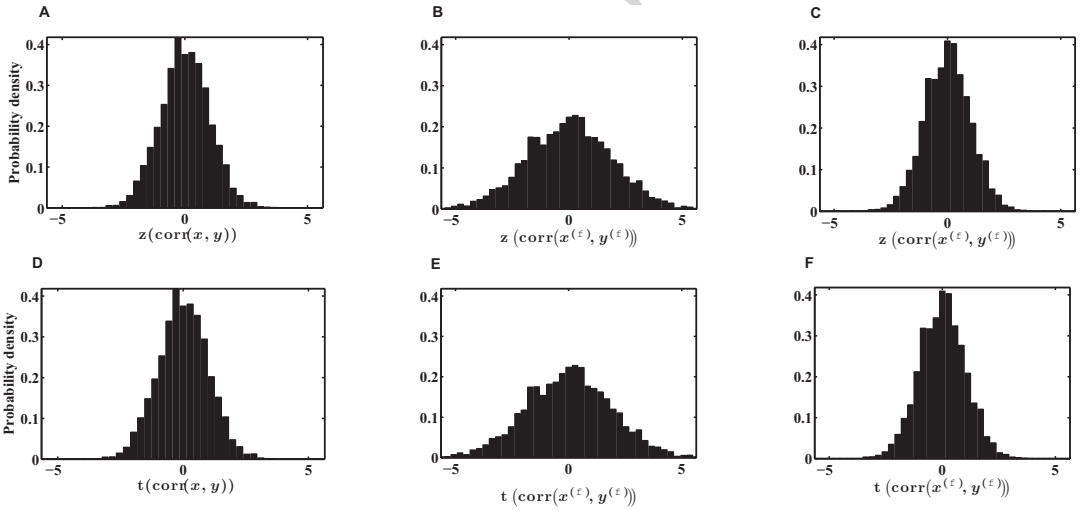

Figure 3: Empirical distributions of correlation test variates derived from Fisher's z-transformation, $\mathrm{z}\left(\mathrm{corr}\left(x,y\right)\right)$, and Student's t-test, $\mathrm{t}\left(\mathrm{corr}\left(x,y\right)\right)$, sampled from both unfiltered and filtered simulation data with; true correlation, $\rho_{x,y} = 0.3$; $T = 1000$; $\sigma^2 = 1$. A low-pass FIR filter with a cut-off frequency of 0.08Hz was used for filtering. **A:** Fisher's z-scores for unfiltered data. **B:** Uncorrected z-scores for filtered data. **C:** Corrected z-scores for filtered data. **D:** t-scores derived from the t-test for unfiltered data. **E:** Uncorrected t-scores for filtered data. **F:** Corrected t-scores for filtered data.

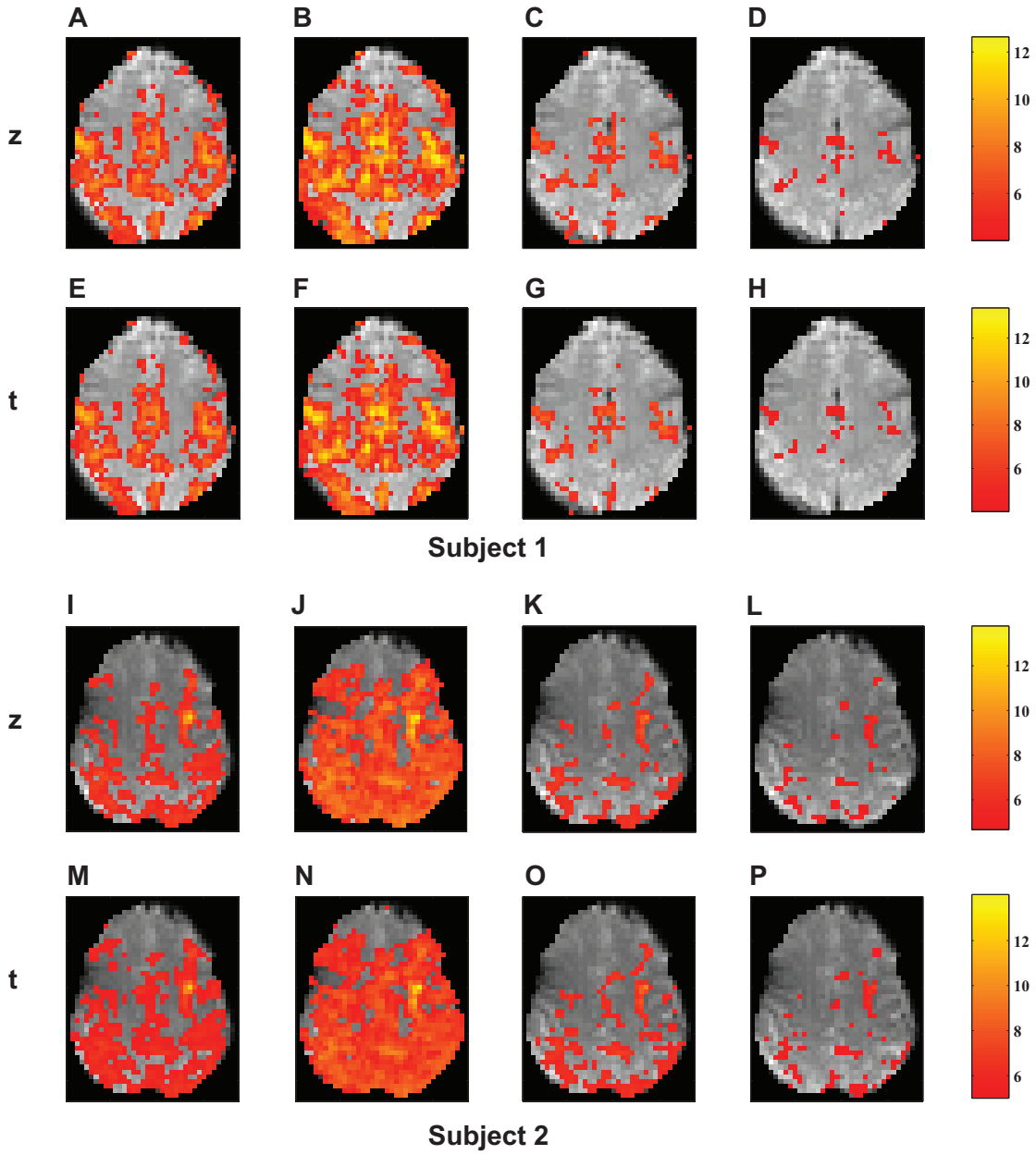

Figure 4: Seed-voxel correlation maps derived from sample correlation values for unfiltered, and filtered, data, significance tested using Fisher's z-transformation, $\mathrm{r}\left(\mathrm{corr}\left(x, y\right)\right)$, and the t-test, $\mathrm{t}\left(\mathrm{corr}\left(x, y\right)\right)$, thresholded at $p < 0.05$ (corrected). An FIR filter with a cut-off frequency of 0.1Hz was used for filtering. Subject 1: panels A-H. Subject 2: panels I-P. **A, I:** Fisher's z-scores for unfiltered data. **B, J:** Uncorrected z-scores for filtered data. **C, K:** Corrected z-scores for filtered data. **D, L:** Corrected z-scores for filtered data modelled to contain intrinsic autocorrelation. **E, M:** t-scores derived from the t-test applied to unfiltered data. **F, N:** Uncorrected t-scores for filtered data. **G, O:** Corrected t-scores for filtered data. **H, P:** Corrected t-scores for filtered data modelled to contain intrinsic autocorrelation.

Filtering induces correlation in fMRI resting state data

Catherine E. Davey, David B. Grayden, Gary F. Egan, and Leigh A. Johnston

Highlights

- We derive the distribution of correlation between filtered timeseries to address sample dependence.
- Corrected distributions are also derived for statistical inference tests of sample correlation.
- Corrected distributions are expressed as a function of the temporal filter's frequency response.
- Proposed corrections are valid for any true correlation and arbitrary filter specifications.
- Empirical and experimental results verify the utility of the proposed corrections.

Author/s:
Davey, CE;Grayden, DB;Egan, GF;Johnston, LA

Title:
Filtering induces correlation in fMRI resting state data

Date:
2013-01-01

Citation:
Davey, C. E., Grayden, D. B., Egan, G. F. & Johnston, L. A. (2013). Filtering induces correlation in fMRI resting state data. NEUROIMAGE, 64 (1), pp.728-740. https://doi.org/10.1016/j.neuroimage.2012.08.022.

Persistent Link:
http://hdl.handle.net/11343/44035