Sliding windows analysis can undo the effects of preprocessing when applied to fMRI data

Martin A. Lindquist
Department of Biostatistics, Johns Hopkins University, Baltimore, MD

October 6, 2023

Abstract

Resting-state fMRI (rs-fMRI) data is used to study the intrinsic functional connectivity (FC) in the human brain. Recently, interest has focused on studying the temporal dynamics of FC on short timescales, ranging from seconds to minutes. These studies of time-varying FC (TVFC) have enabled the classification of whole-brain dynamic FC profiles into distinct "brain states", defined as recurring whole-brain connectivity profiles reliably observed across subjects and sessions. The analysis of rs-fMRI data is complicated by the fact that the measured BOLD signal consists of changes induced by neuronal activation, as well as non-neuronal nuisance fluctuations that should be removed prior to further analysis. Thus, the data undergoes significant preprocesing prior to analysis. In previous work [19], we illustrated the potential pitfalls involved with using modular preprocessing pipelines, showing how later preprocessing steps can *reintroduce* signal variation previously removed from the data. Here we show that the problem runs deeper, and that certain statistical analysis techniques can similarly interact with preprocessing and reintroduce previously removed signal variation. One such technique is the popular sliding window analysis, used to compute TVFC. In this paper, we discuss the problem both theoretically and empirically in application to test-retest rs-fMRI data. Importantly, we show that we are able to obtain essentially the same brain states and state transitions when analyzing motion induced signal as we do when analyzing the preprocessed but windowed data. Our results cast doubt on whether the estimated brain states obtained using sliding window analysis are neuronal in nature, or simply reflect non-neuronal nuisance signal variation (e.g., motion). We highlight the need to critically revisit previous work on rs-fMRI data work that may not have adequately controlled for these types of effects.

Keywords: resting-state fMRI; time-varying connectivity; dynamic connectivity; artifacts; motion; preprocessing

1 Introduction

Resting-state fMRI (rs-fMRI) is used to study the intrinsic functional connectivity (FC) in the human brain ([5]). It has been used to show that fluctuations in the blood oxygen level dependent (BOLD) signal in spatially distant brain regions are strongly and consistently correlated with one another ([3, 11]). Originally, FC was thought to be constant across a given experimental run (i.e., between 5-15 minutes long). However, in recent years, interest has extended to studying the temporal dynamics of FC on much shorter timescales (i.e., seconds to minutes) ([15, 23, 20]). These studies of time-varying FC (TVFC), also known as dynamic connectivity, have enabled the classification of whole-brain dynamic FC profiles into distinct "brain states", that are defined as recurring whole-brain connectivity profiles reliably observed across subjects throughout the course of a resting state run ([8]).

The analysis of rs-fMRI data is complicated by the fact that the measured BOLD signal consists of changes induced by neuronal activation (the signal of interest) and non-neuronal variation. Examples of the latter include drift, spiking artifacts, motion-related artifacts, and fluctuations due to physiological sources (e.g., heart rate and respiration) ([22, 28, 4, 26]). Failure to properly control and correct for these types of variation can have significant impact on subsequent analysis, as they can potentially induce spurious FC between different brain regions. Thus, it is of great interest to remove or at least reduce their effects on the analysis of rs-fMRI data.

For these reasons, rs-fMRI data are subjected to a series of preprocessing steps ([7, 10, 29]) prior to statistical analysis. Many preprocessing pipelines used in the field are modular in nature, by which we mean that they are composed of a number of separate distinct filtering/regression steps. These include, among other steps, removal of head motion covariates using linear regression and band-pass filtering, performed sequentially and in a flexible order. In previous work ([19]), we illustrated the potential pitfalls involved with using modular preprocessing pipelines for data cleaning. We showed how later preprocessing steps can *reintroduce* signal variation removed from the data in prior preprocessing steps. In particular, we illustrated how performing high-pass filtering after motion regression can bring motion induced variation back into the data. Similarly, performing motion regression after high-pass filtering can bring unwanted frequency components back into the signal; a finding that had previously been pointed out by [14].

Mathematically, both of these issues are due to the fact that each regression/filtering step acts as a geometric projection of the data onto a subspace orthogonal to the space spanned by the nuisance covariates/frequencies one seeks to remove, and that performing a sequence of such projections can move the data back into nuisance-related subspaces previously accounted for, thus reintroducing signal related to previously cointrolled nuisance covariates. We showed how these issues can arise in practice when any combination of standard preprocessing steps including motion regression, scrubbing, component-based correction, physiological correction, global signal regression, and temporal fil-

tering are performed sequentially. We also suggested that the problem can be rectified by performing all operations jointly, or alternatively by orthogonalizing all operations with respect to one another.

However, it is important to stress that it is not only preprocessing steps that can interact with one another in such a manner, but perhaps more insidiously statistical analysis techniques can similarly interact with previous preprocessing steps and reintroduce signal variation related to previously removed signal. For example, a statistical analysis technique that incorporates temporal filtering may bring back motion induced signal variation in an analogous manner as outlined above. One such technique commonly used in the analysis of rs-fMRI data is the so-called sliding windows analysis, used to determine TVFC. Here a series of correlation matrices are computed over fixed-length, windowed segments of the fMRI time series. A sliding window is a convolution of signal in the time-domain with a rectangular filter. As seen by the Fourier convolution theorem this is equivalent to applying a low-pass filter in the frequency domain to the data ([18]). Hence, using the same reasoning outlined in [19] it is possible that this operation is inadvertently reintroducing signal variation (for example related to motion induced spin history artifacts ([13])) previously removed during preprocessing back into the signal. Thus, changes in this non-neuronal signal variation may be subsequently mistaken for time-varying connectivity.

In this paper we investigate this issue in detail. We begin by outlining the theoretical concerns, and show how it is possible to reintroduce signal variation previously removed from the data during preprocessing. We thereafter illustrate the problem empirically using test-retest rs-fMRI data. We utilize a standard processing pipeline used to determine brain states in a TVFC study. We illustrate how motion induced signal is reintroduced into the data when using sliding windows and that the subsequent derived brain states are corrupted by motion induced signal variation previously removed during preprocessing. In addition, we show that we are able to obtain essentially the same brain states and transitions when analyzing motion-related signal as we do when analyzing the preprocessed but windowed rs-fMRI data. These results cast doubt on whether the estimated brain states obtained using sliding windows analysis are neuronal in nature, or simply reflect non-neuronal nuisance signal variation. We highlight the need to critically revisit previous work on rs-fMRI data that may not have adequately controlled for these types of effects.

2 Methods

2.1 Computing time-varying functional connectivity

Typically, TVFC between multiple regions of the brain are represented using a series of either covariance or correlation matrices, that depict the relationship between different brain regions or components over time. The most common approach for estimating the elements of the correlation matrix is to use the sliding windows approach. Here, a time window of fixed length w is selected, and

data points within that window are used to calculate the correlation coefficients. The window is then shifted across time and a new correlation coefficient is computed at each time point.

TVFC can be computed using data from individual voxels ([15, 18]), averaged over pre-specified regions of interest ([9]), or estimated using data-driven methods such as independent component analysis ([1]). Once the set of time-varying correlations have been estimated, a standard approach for summarizing them is to classify them into distinct brain states, or recurring whole-brain patterns of FC that appear repeatedly across time and subjects ([1]). To derive brain state, the time-varying correlations across subjects are clustered, typically using k-means ([1]). Practically, this is done by reorganizing the lower triangular portion of each subject's $d \times d \times T$ time-varying correlation data into a matrix with dimensions $((d-1)d/2) \times T$, where d is the number of voxels/regions/components and T the number of time points. The data from all subjects is then concatenated into a matrix with dimensions $((d-1)d/2) \times TN$, where N is the number of subjects. K-means clustering is performed and each of the resulting cluster centroids represent a recurring brain state. The number of clusters can be chosen based on computing the within-group sum of squares for each candidate number of clusters and picking the 'elbow' in the plot; though alternative approaches exist.

2.2 Sliding window as a low-pass filter

Central to our argument is the fact that the convolution of a signal in the time domain with a fixed window acts as a low pass filter in the frequency domain ([18]). To illustrate this fact consider the sliding-window covariance between two time series x and y. The repeated calculation of the local covariance obtained by shifting a window h across time can be written:

$$c_{xy} = (xy) * h - (x * h)(y * h) \tag{1}$$

where $*$ denotes the convolution operator. Using the Fourier convolution theorem, this can be expressed in the frequency domain as:

$$C_{xy} = (X * Y)H - (XH) * (YH) \tag{2}$$

where capital letters represent the DFT of the respective signals. In particular, note that the DFT of a rectangular window is given by the discrete version of a sinc-function, whose main lobe has width $1/w$; see Fig. 1 for an example.

Hence, convolution with a sliding windows in the time-domain is equivalent to applying a low-pass filter in the frequency domain and has similar effects. We therefore hypothesize that this operation will negatively interact with signal removed in previous preprocessing steps in an analogous manner outlined in [19]. While we illustrate the issue using a sliding window covariance for simplicity, these concerns carry over to sliding-window correlations which are normalized functions of the covariance.

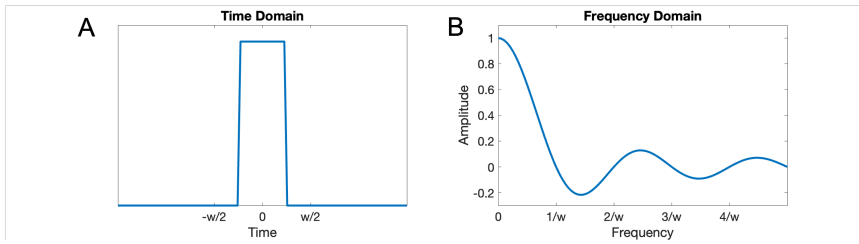

Figure 1: (A) A sliding window of width w represented in the time domain. (B) The discrete Fourier transform of the sliding window takes the form of a sinc function whose main lobe has width $1/w$ in the frequency domain. The sinc function acts as a low-pass filter in the frequency domain.

2.3 Interaction between a sliding window and nuisance regression

We use a geometric framework to illustrate how a sliding window (in the form of a low-pass filter) can interact with previous nuisance regression. We consider both nuisance regression and filtering as linear regressions, and use geometric projections to illustrate our ideas.

Let $\mathbf{y}$ be an n-dimensional vector consisting of rs-fMRI signal from a single voxel in the brain. Further, let $\mathbf{X}$ be an $n \times p$ design matrix whose columns consist of the nuisance regressors we seek to remove from $\mathbf{y}$. This can be done by fitting the following linear model:

$$\mathbf{y} = \mathbf{X}\beta + \epsilon. \tag{3}$$

The least-squares estimate of β is given by

$$\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}, \tag{4}$$

and the fitted value by:

$$\hat{\mathbf{y}} = \mathbf{X}\hat{\beta} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}. \tag{5}$$

Here $\hat{\mathbf{y}}$ represents the estimated nuisance signal in the voxel and corresponds to signal that we seek to remove from $\mathbf{y}$.

To simplify calculations, we take a projection approach towards linear regression. We define a projection matrix

$$\mathbf{H} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' \tag{6}$$

and write $\hat{\mathbf{y}} = \mathbf{H}\mathbf{y}$. Here $\mathbf{H}$ is a linear matrtix transformation that projects $\mathbf{y}$ onto the space spanned by the columns of the design matrix $\mathbf{X}$. An important property of projection matrices are that they are both idempotent ($\mathbf{H} = \mathbf{H}^2$) and symmetric ($\mathbf{H} = \mathbf{H}'$).

To remove the effects of motion we subtract $\hat{\mathbf{y}}$ from the data and compute the residual vector $\mathbf{e} = \mathbf{y} - \hat{\mathbf{y}}$. This can alternatively be expressed $\mathbf{e} = (\mathbf{I} - \mathbf{H})\mathbf{y}$, and this is our new signal of interest. Note $\mathbf{I} - \mathbf{H}$ is also a projection matrix which projects the data onto a subspace orthogonal to the columns of the design matrix $\mathbf{X}$. Thus, our signal of interest $\mathbf{e}$ lies in a subspace orthogonal to $\mathbf{X}$. Hence,

$$\langle \mathbf{Hy}, (\mathbf{I} - \mathbf{H})\mathbf{y} \rangle = \mathbf{y}'\mathbf{H}(\mathbf{I} - \mathbf{H})\mathbf{y} = 0 \tag{7}$$

which is a necessary condition for the data to be uncorrelated with the nuisance signal. Here $\langle ., . \rangle$ represents the Euclidean inner product, and the last equality holds since $\mathbf{H} = \mathbf{H}^2$. At this point we have, to the extent possible using linear projections, removed the effects of the nuisance components in $\mathbf{X}$ from the data.

Next, suppose we perform sliding windows on the resulting data, which as mentioned above is equivalent to low-pass filtering the nuisance-corrected data. Filtering can be expressed in a linear model using a series of sine and cosine functions as regressors in a design matrix Z. Fitting the model

$$\mathbf{e} = \mathbf{Z}\gamma + \epsilon \tag{8}$$

projects the data onto a space corresponding to the frequency components one seeks to retain. We can express this operation as $\mathbf{Pe}$, where $\mathbf{P} = \mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'$ represents the projection matrix into this space. Unfortunately, this operation may project the data into a subspace that is no longer orthogonal to the space spanned by $\mathbf{X}$. This would have the effect of partially reintroducing a relationship between the nuisance signal and the signal being studied. This can be noted by observing that

$$\langle \mathbf{Hy}, \mathbf{P}(\mathbf{I} - \mathbf{H})\mathbf{y} \rangle = \mathbf{y}'\mathbf{HP}(\mathbf{I} - \mathbf{H})\mathbf{y} \tag{9}$$

is no longer guaranteed to be zero. Apart from trivial cases (e.g., $\mathbf{P} = \mathbf{0}$, $\mathbf{H} = \mathbf{0}$, $\mathbf{HP} = \mathbf{0}$), a requirement for the data and motion to be uncorrelated is that $\mathbf{PH} = \mathbf{P}$. The holds if the column space of $\mathbf{Z}$ lies in the column space of $\mathbf{X}$. This ensures that the filter keeps the data in the space orthogonal to that spanned by $\mathbf{X}$. This will generally not hold unless the filter is explicitly designed for this purpose.

In Fig. 2 we illustrate the problem graphically. Here a point in n-dimensional space, represented by the green circle, is sequentially projected onto two different subspaces. Let us assume that the point is first projected onto the light-blue subspace. The projection is shown as an orange circle. Next, assume that this point is further projected onto the gray subspace. The new location is shown as a yellow circle. Panel A illustrates a situation where the two subspaces are orthogonal to one another, allowing the yellow circle to simultaneously lie in both subspaces. Panel B illustrates a situation where the subspaces are not orthogonal, and the yellow circle is no longer in the light-blue subspace. In our example, the light-blue subspace is the space orthogonal to the nuisance regressors that one seeks to remove, and the gray subspace the space spanned

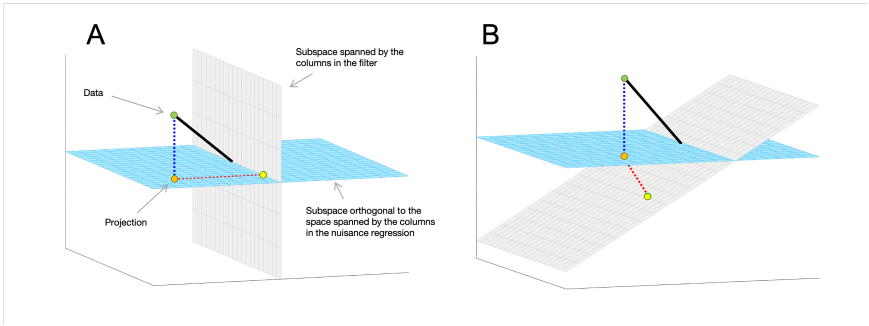

Figure 2: A graphical illustration of the effect of multiple projections. (A) Consider a point in n-dimensional space represented by a green circle. The point is first projected onto the light-blue subspace; see orange circle. This projection is subsequently projected onto the gray subspace; see yellow point. The two subspaces are orthogonal to one another and the second projection simultaneously lies in both spaces. (B) Same setup, but the two subspaces are not orthogonal to one another. Here the second projection is no longer in the light blue subspace.

by the frequency components one seeks to retain. The first projection moves the data into the space orthogonal to the nuisance components, thus removing their influence. Hence, the data, represented by the orange circle, will be uncorrelated with the nuisance regressors. The second projection moves the data into a new subspace. The data, represented by the green circle, lies in the frequency band of interest. If the spaces are orthogonal, as in Panel A, the data are simultaneously uncorrelated with the nuisance regressors and in the frequency band of interest. However, if these two subspaces are not orthogonal as in Panel B, the second projection will move the data back into the space spanned by the nuisance regressors. This has the effect of potentially reintroducing nuisance-induced signal variation into the data.

2.4 Data collection

We used the Multi-Modal MRI Reproducibility Resource (Kirby) from the F.M. Kirby Research Center[1]; see [16] for more information about the acquisition protocol. Briefly, it consists of data from 21 healthy participants scanned on a 3T Philips Achieva scanner. Each participant completed two scanning sessions on the same day, and between sessions a full repositioning of the participant occurred. A T1-weighted MPRAGE structural image was acquired during both sessions (acquisition time $= 6\,min$, TR/TE/TI $= 6.7/3.1/842\,ms$, resolution $= 1 \times 1 \times 1.2\,mm^3$, SENSE factor $= 2$, flip angle $= 8$). A multi-slice SENSE-EPI pulse sequence ([27, 24]) was used to acquire a single rs-fMRI run during each session, consisting of 210 volumes sampled every $2\,s$ at $3\,mm$ isotropic

[1]Publicly available at http://www.nitrc.org/projects/multimodal

spatial resolution (acquisition time: $7\,min$, TE = $30\,ms$, SENSE acceleration factor = 2, flip angle = 75, 37 axial slices collected sequentially with a $1\,mm$ gap). Participants were instructed to rest comfortably while remaining as still as possible, and no other instruction was provided. We refer to the first rs-fMRI session collected as Session 1 and the second as Session 2.

2.5 Preprocessing

SPM 12 (Wellcome Trust Centre for Neuroimaging, London, United Kingdom) and custom MATLAB (The Mathworks, Inc., Natick, MA) was used to preprocess the data. To allow for the stabilization of magnetization, four volumes were discarded at acquisition, and an additional volume was discarded prior to preprocessing. Slice timing correction was performed using the slice acquired at the middle of the TR as reference, and rigid body realignment parameters were estimated to adjust for head motion. The six rigid-body parameters were saved for further analysis; see below. Structural runs were registered to the first functional frame and spatially normalized to Montreal Neurological Institute (MNI) space using SPM's unified segmentation-normalization algorithm ([2]). The estimated rigid body and nonlinear spatial transformations were applied to the rs-fMRI data.

2.6 Creation of nuisance images

Before further preprocessing, we estimated the total amount of motion at each voxel of the brain for each subject and session. To do so we created a design matrix $\mathbf{X}$ consisting of 24 regressors. These included the six motion regressors obtained after rigid-body transformation, the regressors squared, their first order difference, and the first order difference squared. Together these 24 regressors have been shown to appropriately model so-called spin-history artifacts (i.e., residual effects of motion after rigid-body correction) ([13]). The fitted values of this motion regression are typically removed from the signal prior to further analysis, giving rise to what we will refer to as motion-corrected images. Fig. 3 provides a graphical illustration of the approach.

Here $\mathbf{y}_v$ corresponds to the signal of interest from voxel v, and $\mathbf{X}$ the design matrix containing the 24 motion regressors. To remove the effect of motion we fit the motion regression $\mathbf{y}_v = \mathbf{X}\beta_v + \epsilon$. After estimating the parameters β_v corresponding to the motion regressors at each voxel v, we use them to create 'nuisance images'. This was done by computing the fitted values $\mathbf{X}\beta_v$ at each voxel, thereby creating 4D images of the estimated motion induced signal variation at each voxel in the brain; see Fig. 3D. These nuisance images can be extended to incorporate other non-neuronal signal (e.g., signal from white matter or cerebral spinal fluid), but we only consider motion here for simplicity. These images provide a summary of the information related to nuisance covariates removed from the signal.

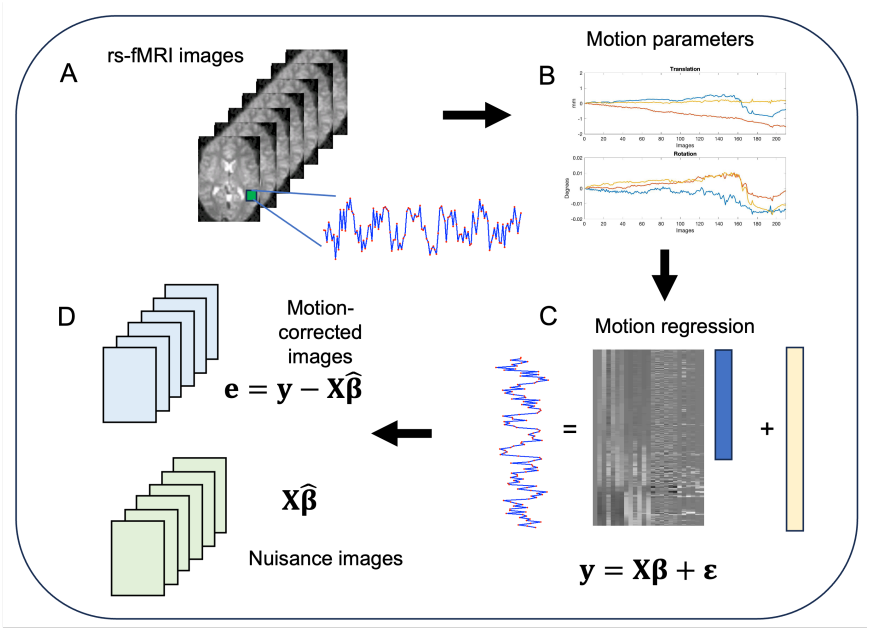

Figure 3: An illustration of the procedure to create nuisance images. (A) A signal is extracted from a single voxel over time. (B) Prior to extraction of the signal a rigid body realignment is performed to adjust for head motion, and the six parameters were saved for further analysis. (C) The six motion parameters obtained after rigid-body transformation, the parameters squared, their first order difference, and the first order difference squared are used as predictors in a linear model to remove additional motion-related artifacts in a process known as motion regression. (D) In the motion-corrected images, the effects of motion are removed from the signal of interest and used to create new data orthogonal to this confounder. In the nuisance images the fitted values from the motion regression are used to create new data consisting purely of nuisance signal.

2.7 Analysis pipelines

Next we applied four different analytic pipelines to the data, outlined below.

Pipeline 1: In the first pipeline, we first perform motion regression (as described above) and remove the estimated motion from the rs-fMRI data. We thereafter parcellated the data into 268 non-underlapping regions using the Shen atlas ([25]), and performed sliding windows correlation and k-means clustering. We used a window length of 30 TRs for the window, which is the suggested optimal window-length for rs-fMRI data collected using a standard EPI sequence with a sampling frequency of 2s ([18]). K-means clustering was repeated 50 times, using random initialization of centroid positions, to increase the chance of escaping local minima. The elbow method was used to choose the two cluster solution, corresponding to two reoccurring brain states. This was repeated for both sessions for each of the 21 subjects.

Pipeline 2: In the second pipeline, we repeated the same procedure on rs-fMRI data without first performing motion regression. Hence, in this pipeline the effects of motion should still be present in the signal on which TVFC was assessed.

Pipeline 3: In the third pipeline, we repeated the same procedure described in Pipeline 1 on the nuisance images. Importantly, these images consist entirely of motion induced signal variation that was removed in Pipeline 1. We perform this pipeline to get a clearer picture of the effects of motion on the estimated TVFC.

Pipeline 4: Finally, in the fourth pipeline we used the same data as in Pipeline 1, but perform a secondary motion regression procedure inside of each sliding window. [21] suggested a variant of this approach, referred to as block regression, as a way to mitigate the effects of motion artifacts when using sliding windows. We use this approach as it allows motion regression and low pass filtering (in the form on the sliding window) to be performed simultaneously, so not to interact with one another.

2.8 Assessing the Impact of Motion

To investigate the impact of motion at each step of the analysis we compute the average fitted value from the motion regression over each region (corresponding to the regional averages of the nuisance images). We compute the correlation between this signal and the average rs-fMRI over each region both with and without motion regression. The former corresponds to Pipeline 1 and the latter to Pipeline 2. To investigate whether the application of the sliding window kernel reintroduces motion induced signal variation, we further computed the correlation between the motion signal and the resting-state signal after application of a sliding window kernel. This is meant to illustrate how applying a low-pass filter (i.e., the sliding window) reintroduces motion induced signal variation. This process was repeated after the secondary motion regression was

performed (Pipeline 4).

The states estimated using the four different pipelines were compared to one another using multidimensional scaling (MDS). MDS provides a distance-preserving map of the input data in two-dimensional space, thus providing a visual representation of the distance between the different brain states. This allows us to evaluate 'how close' brain states obtained using the various pipelines lie to one another.

Finally, to assess the impact of motion on Pipeline 1, which is a commonly used pipeline, we estimated the proportion motion induced signal variation after windowing using the following metric: $R_1^2/(R_1^2 + R_2^2)$. Here R_1^2 is the squared correlation between the motion data and the windowed time courses expressed as $corr(\mathbf{Hy}, \mathbf{P}(\mathbf{I}-\mathbf{H})\mathbf{y})$ and R_2^2 is the squared correlation between the motion-corrected data and the windowed time courses expressed as $corr((\mathbf{I}-\mathbf{H})\mathbf{y}, \mathbf{P}(\mathbf{I}-\mathbf{H})\mathbf{y})$. This metric allows to to investigate the proportion of motion induced signal variation present in the signal after the windowing procedure. Prior to windowing, this metric will be 0 in all regions of the brain.

3 Results

Fig. 4A shows the correlation between the signal and motion in each region using the data where no motion regression was preformed (Pipeline 2). Clearly, the signal is highly correlated with motion, showing that motion effects play a prominent role in the signal without additional preprocessing. Panel B shows the same correlation in the data after motion regression was preformed (Pipeline 1). In this setting, as expected all motion has been properly removed from the signal. However, Panel C shows the correlation between the windowed motion corrected time course and the motion. Here the windowing process has lead to a reintroduction of motion induced signal variation and motion effects are once more correlated with the signal of interest. Finally, Panel D shows the results between the windowed time course and the motion, in the setting where motion regression is performed within each window. This procedure has led to a significant reduction in correlation, though not a complete removal; see Section 3 for further discussion.

Fig. 5 shows the estimated brain states (left) and transition times (right) for each of the four pipelines. Starting with Pipeline 3 (third row), we see the states estimated using the nuisance images. The first state shows that most brain regions are highly interconnected, while the second state is significantly less connected. These states are further split into 8 separate functional networks as described in [12]: medial frontal (purple), frontoparietal (cyan), default mode (yellow), subcortical-cerebellum (red), motor (blue), visual I (pink), visual II (green), and visual association (maroon).

Interestingly, the derived states from Pipeline 3 are similar to those found using Pipelines 1 and 2. These are the states where motion was removed and not removed prior to analysis, respectively. In many ways they are not differentiable from the nuisance states. The transition times for the three first

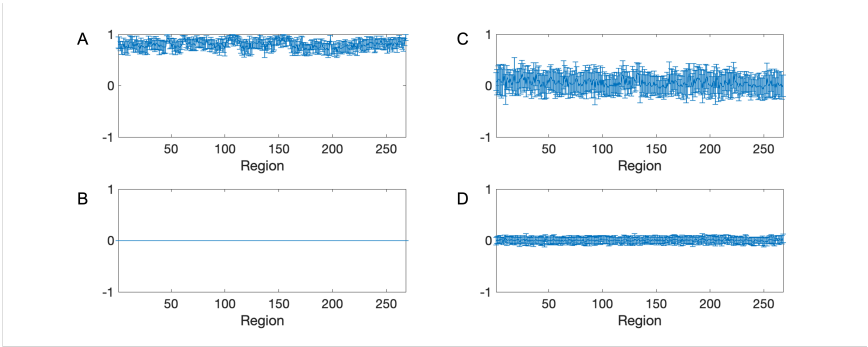

Figure 4: Group-level results for Session 1. (A) The correlation between the motion signal and the non motion-corrected rs-fMRI signal over each region (Pipeline 2). (B) The correlation between the motion signal and the motion-corrected rs-fMRI signal over each region (Pipeline 1). (C) The correlation between the motion signal and the resting-state signal after the application of a sliding window kernel (Pipeline 1); illustrating the reintroduction of motion induced signal variation after applying the sliding window. (D) The correlation between the motion signal and the resting-state signal after the application of a sliding window kernel and secondary motion regression (Pipeline 4).

pipelines are also similar, indicating that the transitions may be highly related to motion induced signal variation. The results from Pipeline 4 (fourth row) appears different. The correlations between regions is generally decreased and a number of cross-correlations between brain regions in State 2 are less apparent, changing the interpretation of the findings. Finally, the transition times are quite different from those seem in the other pipelines, showing more frequent transitions between states.

The similarity between Pipelines is further investigated in Fig. 6A. Here multidimensional scaling was applied to the eight States; two States derived from each of the four Pipelines. The placement of the states in the first two dimensions are shown. Circles correspond to State 1 and squares to State 2, while the different pipelines are color-coded according to the figure legend. The green symbols correspond to the nuisance states (Pipeline 3). The States derived from Pipeline 2 (red symbols) are close to them, while the States obtained after performing an initial motion correction (Pipeline 1; blue symbols) lie yet further away. Finally, performing an additional motion regression moves the States yet further away from the nuisance states (Pipeline 4; cyan symbols). Panel C shows the proportion nuisance present in the signal given by $R_1^2/(R_1^2+R_2^2)$. Here R_1^2 is the squared correlation between the motion data and the windowed time courses and R_2^2 is the squared correlation between the motion-corrected data and the windowed time courses. The results indicate a substantial amount of motion induced signal variation is present. Fig. S1 shows the same proportion averaged over each of the 8 functional networks. In particular, the subcortical-cerebellum

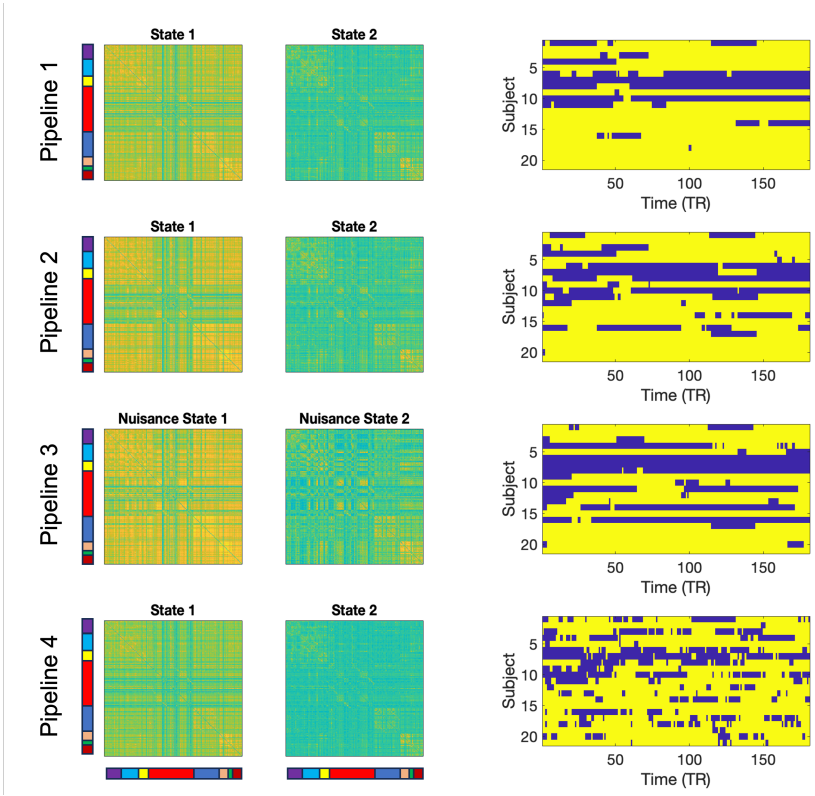

Figure 5: Results for Session 1. The two estimated brain states (left) and transition times between states (right) for each of the four pipelines. Each pipeline gives rise to two states, and the transitions between states across time is concatenated across subjects giving a total of 180×21 time points.

and motor networks show significant impact of motion induced variation, with values ranging from 0.25 to 0.45. Finally, Panel D shows a heatmap of the frame-wise displacement measured at every time point for every subject, illustrating the amount of motion for each subject.

Importantly, these results are reproducible using data obtained on the same subjects from a different scanning session. The equivalent results are seen in Figs. 7, 8 and 6B, E and F, which show strong similarity with the findings from the first session.

More granular results for a single randomly chosen subject (ID number 02) is seen in Supplementary Figs. S2 and S3. It reinforces many of the findings seen in the group-level data. Importantly, the similarity in transition times between the first three pipelines is apparent in Supplementary Fig. S3.

4 Discussion

In this paper we illustrate how sliding windows, a commonly used technique for assessing time-varying functional connectivity of rs-fMRI data, can interact with previous data preprocessing steps, thereby potentially reintroducing previously removed signal variation. Sliding windows analysis entails convolving a signal in the time-domain with a rectangular (or tapered) filter. As seen by the Fourier convolution theorem this is equivalent to applying a low-pass filter in the frequency domain to the data. We show that this process inadvertently reintroduces motion induced signal variation, that was previously removed in the data cleaning component of the analysis, and that these motion effects can potentially be mistaken for time-varying connectivity.

There has long been a debate over whether nuisance effects may be accounting for a significant portion of the observed fluctuations in TVFC. For example, [17] studied the effects of high-motion frames and came to the conclusion that much of the observed dynamics were related to motion. [15] proposed that nuisance effects may be have larger impact when using a sliding-window approach due to the fact that transient effects can alter correlation estimates when using short windows. Finally, [21] showed that sliding windows correlation estimates can be strongly and significantly correlated with the sliding window norms associated with various nuisance measures. They further found that the correlation can persist even after nuisance regression has been performed. They hypothesized several potential reasons for these observations, including the possibility that a large fraction of the nuisance term is orthogonal to the subspace spanned by the BOLD signals being studied. In this work we find equivalent results, but illustrate that it is caused by the interaction between the sliding window kernel and the nuisance regression. To the best of our knowledge this link has not been noted in previous work.

This work complements previous work in the field ([14, 6, 19]) that showed how different preprocessing steps can interact with one another and thus counteract each others effect. The current work shows that perhaps more insidiously, later statistical analysis can similarly counteract previously performed prepro-

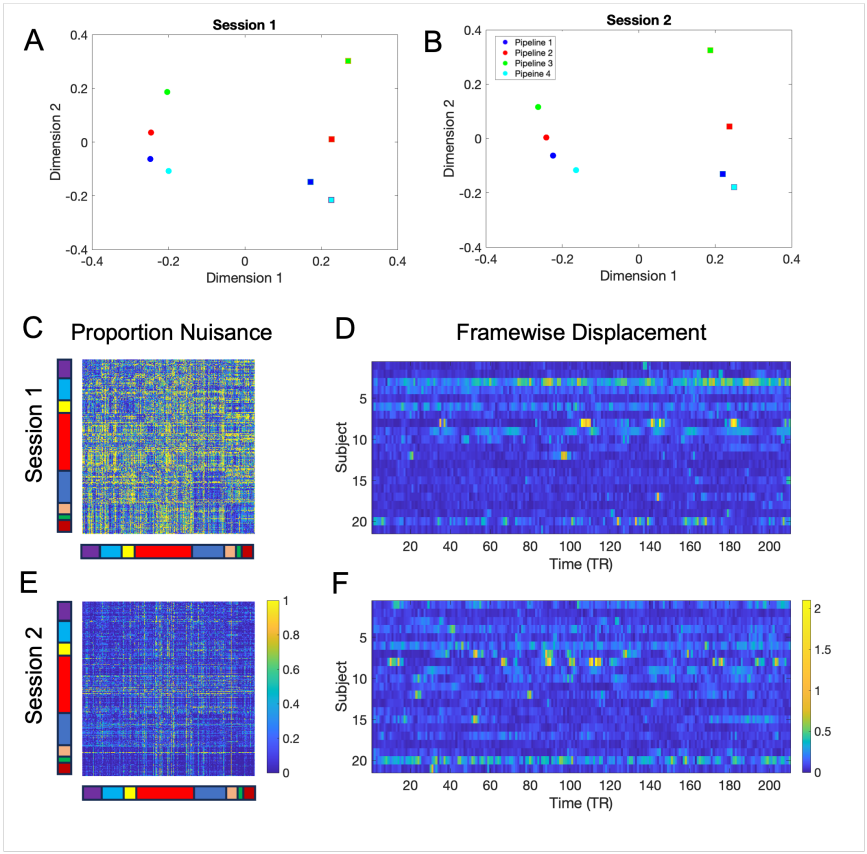

Figure 6: Results for Sessions 1 and 2. (A-B) Multidimensional scaling applied to the two States derived for each of the four Pipelines. Circles correspond to State 1 and squares to State 2. The different pipelines are color-coded according to the figure legend. (A) Shows the results for Session 1 and (B) shows the results for Session 2. (C) The proportion nuisance $R_1^2/(R_1^2 + R_2^2)$, where R_1^2 is the squared correlation between the motion data and the windowed time courses and R_2^2 is the squared correlation between the motion-corrected data and the windowed time courses. (D) A heatmap of the frame-wise displacement measured at every time point for every subject. (E) Shows the same results as (C) for Session 2. (F) Shows the same results as (D) for Session 2.

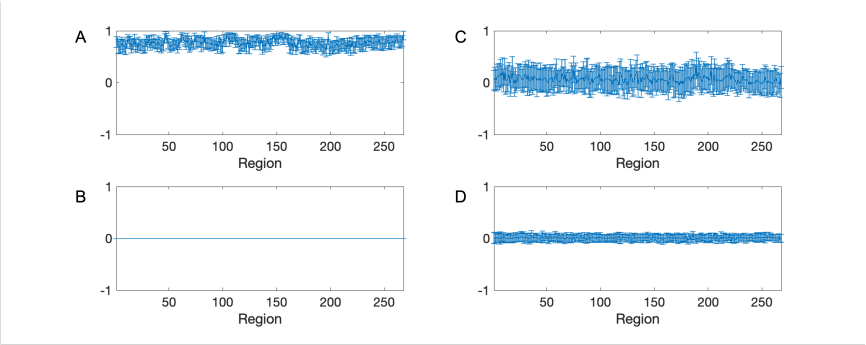

Figure 7: Group-level results for Session 2. (A) The correlation between the motion signal and the non motion-corrected rs-fMRI signal over each region (Pipeline 2). (B) The correlation between the motion signal and the motion-corrected rs-fMRI signal over each region (Pipeline 1). (C) The correlation between the motion signal and the resting-state signal after the application of a sliding window kernel (Pipeline 1); illustrating the reintroduction of noise after applying the sliding window. (D) The correlation between the motion signal and the resting-state signal after the application of a sliding window kernel and secondary motion regression (Pipeline 4).

cessing. We illustrate using a sliding window analysis, but the same problem should occur in any technique utilizing a moving window to assess TVFC.

This type of interaction is particularly problematic as the different components of data analysis have become increasingly compartmentalized. For example, in many research projects one group may acquire the data and do initial processing to place the data into a certain format, another group may preprocess or clean the data, while a third group may perform statistical analysis to the processed data. This compartmentalization has the effect of binning researchers into different non-overlapping silos, that may or may not actually communicate with one another even when working on the same dataset. This bifurcation makes it more difficult to safeguard against the issues raised in this paper.

How do we guard against such problems? We offer several suggestions. First, we find it helpful to run nuisance signal through the analysis pipeline and compare results with those obtained with the preprocessed data; see our Pipeline 3 for an example. This will allow researchers to study the similarity between findings to determine potential interactions. Second, it is important to train researchers to understand the entire analytic pipeline from acquisition to final analysis. Ideally, the entire pipeline should be run together, with a careful examination of interactions between parts of the pipeline. If interactions are detected, these parts of the pipeline can be run simultaneously (e.g., as in Pipeline 4) or the different components can be orthogonalized with respect to one another. It is critical for researchers to critically evaluate commonly used analytic pipeline for flaws. We stress that these can sometimes be quite

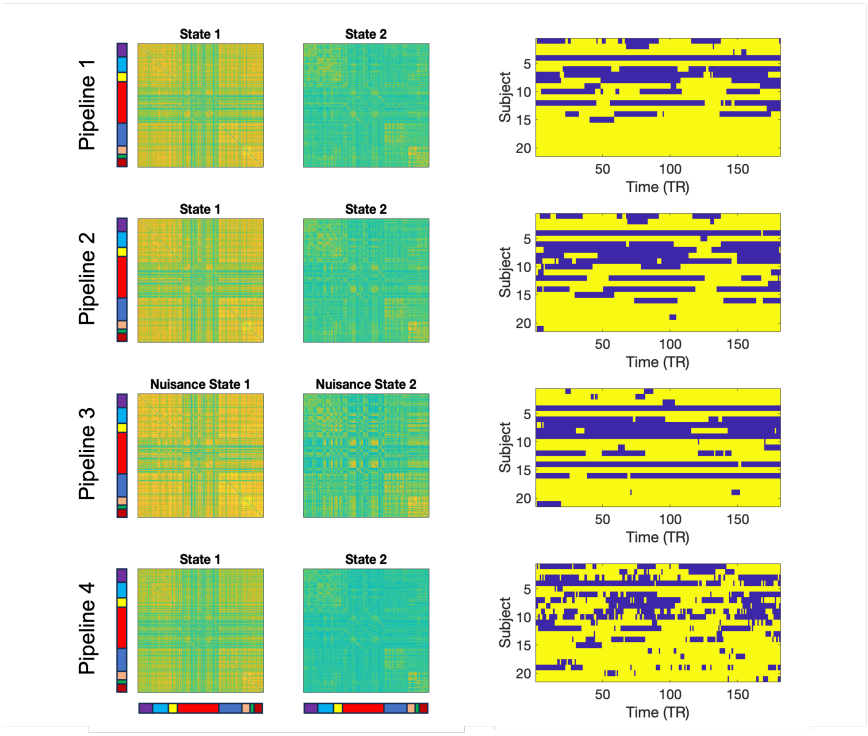

Figure 8: Results for Session 2. The two estimated brain states (left) and transition times between states (right) for each of the four pipelines. Each pipeline gives rise to two states, and the transitions between states across time is concatenated across subjects giving a total of 180×21 time points.

subtle and occur in unexpected ways. Further, it should be possible to develop techniques for removing motion that do not interact with the filtering procedure. Our Pipeline 4 is a simple approach that does most of the job. However, it does not completely remove the effects of motion. This is partially due to the fact that the process was performed on the region-averaged signal rather than on the individual voxels for computational efficiency. In addition, the amount of data within a window makes performing motion regression in this manner difficult. In our illustration, we used a window length of 30 TR. If we perform motion regression consisting of 24 motion regressors this quickly uses all of the degrees of freedom available in the data. If we further seek to include additional covariates (e.g., CSF, white matter signal, global signal regression), we will have more regressors than observations. It may ultimately, be best to create an omnibus analysis technique that performs preprocessing and time-varying connectivity jointly. However, we leave the development of a more comprehensive approach to future work.

A natural question to ask is whether one can avoid these problems by choosing an appropriate window length for the applied sliding-window. It may be possible to choose a window length that gives rise to a filter that minimally interacts with a certain type of nuisance signal. However, though we illustrated the problem using motion induced signal, we stress these issues can occur for any potential nuisance covariate, including scrubbing, component-based correction, physiological correction, and global signal regression. We anticipate that it would be difficult to safeguard against reintroduction of all of these artifacts by choosing a single 'optimal' window length.

How serious is this problem in practice? It is of course difficult to make a blanket statement on this issues. Fig. S1 provides an estimate of the proportion nuisance induced signal variation present in the signal averaged over 8 functional networks. In particular, the subcortical-cerebellum and motor networks show significant impact of motion induced signal variation, with values ranging from 0.25 to 0.45 in Session 1. The proportion was lower, but still non-trivial in Session 2, with values ranging from 0.15 to 0.25.

Importantly, in our example data set, the nuisance data and the 'cleaned' data give rise to similar state estimates and state transitions. This is concerning. It is naturally reasonable to question how 'pure' each of these signals are in practice. For example, the motion-corrected data presumably still includes some degree of motion induced signal, and similarly the nuisance images contain some neuronal signal. If these components cannot be adequately separated it may be the case that similarities in the state estimation is driven by this contamination. Regardless, we believe it is clear that the ultimate goal of preprocessing is to remove as much of these nuisance-induced signal variation as possible, after all otherwise why perform preprocessing in the first place? If so, it should not be considered acceptable to apply a statistical technique that potentially negates this effort. As a field we should seek to develop better methods that don't allow for this possibility.

In summary, this work highlights the potential for the effects of confounding variables being reintroduced into the data when sliding windows analysis is

performed. This may negate the assumption that these effects are adequately controlled in preprocessing. The results raise the questions to what degree estimated brain states reflect non-neuronal nuisance signal variation (e.g., related to motion) rather than the neuronal signal of interest.

Acknowledgments

The work presented in this paper was supported in part by NIH grants R01 EB026549 from the National Institute of Biomedical Imaging and Bioengineering and R01 MH129397 from the National Institute of Mental Health. The author thanks Brian Caffo, Ciprian Crainiceanu, and Jim Pekar for helpful comments on the manuscript.

References

[1] E. A. Allen, E. Damaraju, S. M. Plis, E. B. Erhardt, T. Eichele, and V. D. Calhoun. Tracking whole-brain connectivity dynamics in the resting state. *Cerebral cortex*, 24(3):663–676, 2014.

[2] J. Ashburner and K. J. Friston. Unified segmentation. *Neuroimage*, 26(3):839–851, 2005.

[3] C. F. Beckmann, M. DeLuca, J. T. Devlin, and S. M. Smith. Investigations into resting-state connectivity using independent component analysis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360(1457):1001–1013, 2005.

[4] R. M. Birn, J. B. Diamond, M. A. Smith, and P. A. Bandettini. Separating respiratory-variation-related fluctuations from neuronal-activity-related fluctuations in fmri. *Neuroimage*, 31(4):1536–1548, 2006.

[5] B. Biswal, F. Zerrin Yetkin, V. M. Haughton, and J. S. Hyde. Functional connectivity in the motor cortex of resting human brain using echo-planar mri. *Magnetic resonance in medicine*, 34(4):537–541, 1995.

[6] M. G. Bright, C. R. Tench, and K. Murphy. Potential pitfalls when denoising resting state fmri data using nuisance regression. *Neuroimage*, 154:159–168, 2017.

[7] C. Caballero-Gaudes and R. C. Reynolds. Methods for cleaning the bold fmri signal. *Neuroimage*, 154:128–149, 2017.

[8] V. D. Calhoun, R. Miller, G. Pearlson, and T. Adalı. The chronnectome: time-varying connectivity networks as the next frontier in fmri data discovery. *Neuron*, 84(2):262–274, 2014.

[9] C. Chang and G. H. Glover. Time–frequency dynamics of resting-state brain connectivity measured with fmri. *Neuroimage*, 50(1):81–98, 2010.

[10] R. Ciric, D. H. Wolf, J. D. Power, D. R. Roalf, G. L. Baum, K. Ruparel, R. T. Shinohara, M. A. Elliott, S. B. Eickhoff, C. Davatzikos, et al. Benchmarking of participant-level confound regression strategies for the control of motion artifact in studies of functional connectivity. *Neuroimage*, 154:174–187, 2017.

[11] M. De Luca, C. Beckmann, N. De Stefano, P. Matthews, and S. M. Smith. fmri resting state networks define distinct modes of long-distance interactions in the human brain. *Neuroimage*, 29(4):1359–1367, 2006.

[12] E. S. Finn, X. Shen, D. Scheinost, M. D. Rosenberg, J. Huang, M. M. Chun, X. Papademetris, and R. T. Constable. Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. *Nature neuroscience*, 18(11):1664, 2015.

[13] K. J. Friston, S. Williams, R. Howard, R. S. Frackowiak, and R. Turner. Movement-related effects in fmri time-series. *Magnetic resonance in medicine*, 35(3):346–355, 1996.

[14] M. N. Hallquist, K. Hwang, and B. Luna. The nuisance of nuisance regression: spectral misspecification in a common approach to resting-state fmri preprocessing reintroduces noise and obscures functional connectivity. *Neuroimage*, 82:208–225, 2013.

[15] R. M. Hutchison, T. Womelsdorf, E. A. Allen, P. A. Bandettini, V. D. Calhoun, M. Corbetta, S. Della Penna, J. H. Duyn, G. H. Glover, J. Gonzalez-Castillo, et al. Dynamic functional connectivity: promise, issues, and interpretations. *Neuroimage*, 80:360–378, 2013.

[16] B. A. Landman, A. J. Huang, A. Gifford, D. S. Vikram, I. A. L. Lim, J. A. Farrell, J. A. Bogovic, J. Hua, M. Chen, S. Jarso, et al. Multi-parametric neuroimaging reproducibility: a 3-t resource study. *Neuroimage*, 54(4):2854–2866, 2011.

[17] T. O. Laumann, A. Z. Snyder, A. Mitra, E. M. Gordon, C. Gratton, B. Adeyemo, A. W. Gilmore, S. M. Nelson, J. J. Berg, D. J. Greene, et al. On the stability of bold fmri correlations. *Cerebral cortex*, 27(10):4719–4732, 2017.

[18] N. Leonardi and D. Van De Ville. On spurious and real fluctuations of dynamic functional connectivity during rest. *Neuroimage*, 104:430–436, 2015.

[19] M. A. Lindquist, S. Geuter, T. D. Wager, and B. S. Caffo. Modular preprocessing pipelines can reintroduce artifacts into fmri data. *Human brain mapping*, 40(8):2358–2376, 2019.

[20] D. J. Lurie, D. Kessler, D. S. Bassett, R. F. Betzel, M. Breakspear, S. Kheilholz, A. Kucyi, R. Liégeois, M. A. Lindquist, A. R. McIntosh, et al. Questions and controversies in the study of time-varying functional connectivity in resting fmri. *Network neuroscience*, 4(1):30–69, 2020.

[21] A. Nalci, B. D. Rao, and T. T. Liu. Nuisance effects and the limitations of nuisance regression in dynamic functional connectivity fmri. *Neuroimage*, 184:1005–1031, 2019.

[22] J. D. Power, A. Mitra, T. O. Laumann, A. Z. Snyder, B. L. Schlaggar, and S. E. Petersen. Methods to detect, characterize, and remove motion artifact in resting state fmri. *Neuroimage*, 84:320–341, 2014.

[23] M. G. Preti, T. A. Bolton, and D. Van De Ville. The dynamic functional connectome: State-of-the-art and perspectives. *Neuroimage*, 160:41–54, 2017.

[24] K. P. Pruessmann, M. Weiger, M. B. Scheidegger, P. Boesiger, et al. Sense: sensitivity encoding for fast mri. *Magnetic resonance in medicine*, 42(5):952–962, 1999.

[25] X. Shen, F. Tokoglu, X. Papademetris, and R. T. Constable. Groupwise whole-brain parcellation from resting-state fmri data for network node identification. *Neuroimage*, 82:403–415, 2013.

[26] K. Shmueli, P. van Gelderen, J. A. de Zwart, S. G. Horovitz, M. Fukunaga, J. M. Jansma, and J. H. Duyn. Low-frequency fluctuations in the cardiac rate as a source of variance in the resting-state fmri bold signal. *Neuroimage*, 38(2):306–320, 2007.

[27] M. K. Stehling, R. Turner, and P. Mansfield. Echo-planar imaging: magnetic resonance imaging in a fraction of a second. *Science*, 254(5028):43–50, 1991.

[28] K. R. Van Dijk, M. R. Sabuncu, and R. L. Buckner. The influence of head motion on intrinsic functional connectivity mri. *Neuroimage*, 59(1):431–438, 2012.

[29] C. Yan and Y. Zang. Dparsf: a matlab toolbox for" pipeline" data analysis of resting-state fmri. *Frontiers in systems neuroscience*, 4:13, 2010.

5 Supplementary Figures

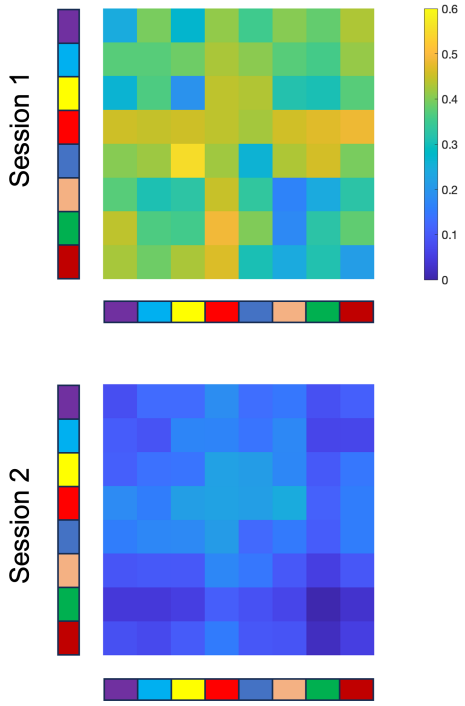

Figure S1: The proportion nuisance $R_1^2/(R_1^2 + R_2^2)$, where R_1^2 is the squared correlation between the motion data and the windowed time courses and R_2^2 is the squared correlation between the motion-corrected data and the windowed time courses averaged over 8 different functional networks. The top panel shows the results for Session 1 and the bottom panel shows the results for Session 2.

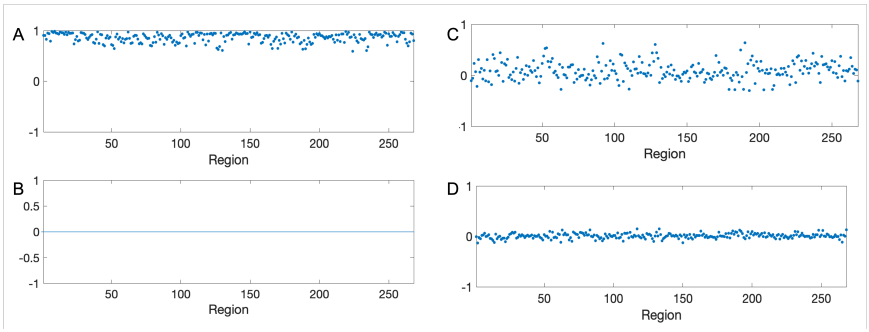

Figure S2: Group-level results for Subject 2. (A) The correlation between the motion signal and the non motion- corrected rs-fMRI signal over each region (Pipeline 3). (B) The correlation between the motion signal and the motion-corrected rs-fMRI signal over each region (Pipeline 1). (C) The correlation between the motion signal and the resting-state signal after the application of a sliding window kernel (Pipeline 1); illustrating the reintroduction of noise after applying the sliding window. (D) The correlation between the motion signal and the resting-state signal after the application of a sliding window kernel and secondary motion regression (Pipeline 4).

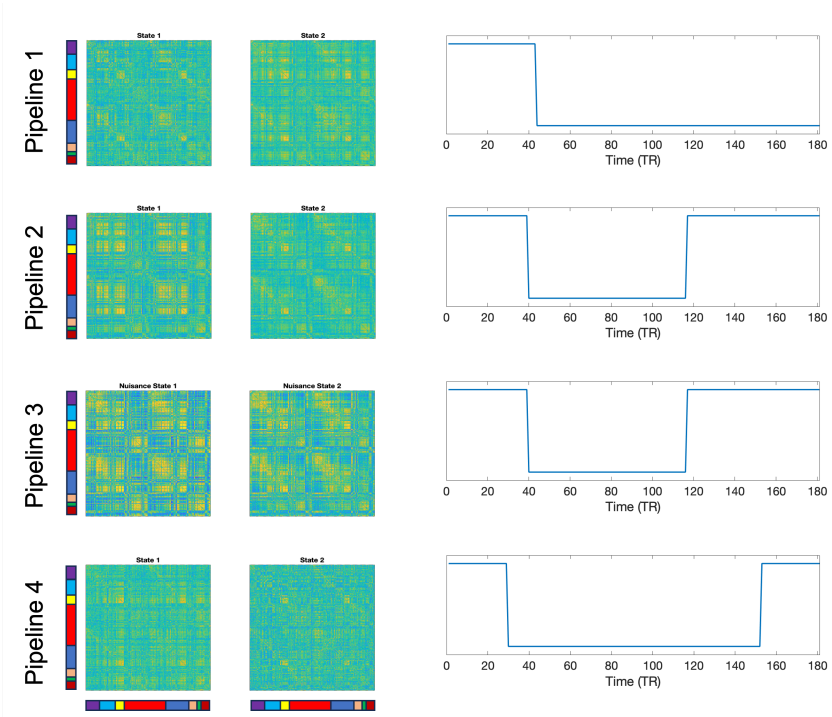

Figure S3: Results for Subject 2. The two estimated brain states (left) and transition times between states (right) for each of the four pipelines. Each pipeline gives rise to two states, and the transitions between states across time is concatenated across subjects giving a total of 180×21 time points.