Single-subject Single-session Temporally-Independent Functional Modes of Brain Activity

Daniel E.P. Gomez a,*, Alberto Llera a, b, José P.R. F. Marques a, Christian F. Beckmann a, b, c, d, David G. Norris a, c, e

a Radboud University Nijmegen, Donders Institute for Brain, Cognition and Behaviour, Donders Centre for Cognitive Neuroimaging, Nijmegen, the Netherlands
b Department of Cognitive Neuroscience, Radboud University Medical Centre, Nijmegen, the Netherlands
c MIRA Institute for Biomedical Technology and Technical Medicine, University of Twente, Enschede, the Netherlands
d Oxford Centre for Functional Magnetic Resonance Imaging of the Brain (FMRIB), University of Oxford, Oxford, OX3 9DU, United Kingdom
e Erwin L. Hahn Institute for Magnetic Resonance Imaging, UNESCO-Weltkulturerbe Zollverein, Essen, Germany

ARTICLE INFO
Keywords:
- multiband
- sms
- mesh
- fMRI
- Ultra-fast
- tfm
- resting-state
- connectivity

ABSTRACT
Temporally independent functional modes (TFMs) are functional brain networks identified based on their temporal independence. The rationale behind identifying TFMs is that different functional networks may share a common anatomical infrastructure yet display distinct temporal dynamics. Extracting TFMs usually require a larger number of samples than acquired in standard fMRI experiments, and thus have therefore previously only been performed at the group level. Here, using an ultra-fast fMRI sequence, MESH-EPI, with a volume repetition time of 158 ms, we conducted an exploratory study with n = 6 subjects and computed TFMs at the single subject level on both task and resting-state datasets. We identified 6 common temporal modes of activity in our participants, including a temporal default mode showing patterns of anti-correlation between the default mode and the task-positive networks, a lateralised motor mode and a visual mode integrating the visual cortex and the visual streams. In alignment with other findings reported recently, we also showed that independent time-series are largely free from confound contamination. In particular for ultra-fast fMRI, TFMs can separate the cardiac signal from other fluctuations. Using a non-linear dimensionality reduction technique, UMAP, we obtained preliminary evidence that combinations of spatial networks as described by the TFM model are highly individual. Our results show that it is feasible to measure reproducible TFMs at the single-subject level, opening new possibilities for investigating functional networks and their integration. Finally, we provide a python toolbox for generating TFMs and comment on possible applications of the technique and avenues for further investigation.

1. Introduction
Resting-state BOLD functional magnetic resonance (fMRI) is a non-invasive neuroimaging technique used to explore the brain’s functional connectivity (FC), as it may relate to cognition, behaviour, mental health and human development (Llera Arenas, Wolfers, Mulders and Beckmann, 2018; Mather et al., 2013; Mulders et al., 2015). FC estimates from resting-state fMRI data are obtained by measuring time-course similarities between different regions of the brain. For this purpose, methods such as seed based correlation analysis and spatial independent component analysis (ICA) (Beckmann and Smith, 2004) are commonly used to identify functional networks, also referred to as resting-state networks (RSNs). Resting State Networks identified from BOLD fluctuations are robust and have been shown to correspond to known functional systems (Smith et al., 2009). Because they may offer insights into cognitive processes, are highly reproducible and easy to measure experimentally, and can be identified with a high temporal and spatial resolution, they have become a powerful tool for neuroimaging.

A limitation of standard FC analysis techniques is that they tend to characterise functional networks as being collections of distinct anatomical regions, either by searching for time-course correlations between voxels or ROIs, or by maximizing spatial independence between different regions of the brain, as in the case of spatial ICA. These techniques are, by design, sub-optimal if one wants to explore distinct

* Corresponding author.
E-mail addresses: d.gomez@donders.ru.nl (D.E.P. Gomez), a.llera@donders.ru.nl (A. Llera), j.marques@donders.ru.nl (J.P.R.F. Marques), c.beckmann@donders.ru.nl (C.F. Beckmann), d.norris@donders.ru.nl (D.G. Norris).

https://doi.org/10.1016/j.neuroimage.2020.116783
Received 23 April 2019; Received in revised form 10 February 2020; Accepted 3 April 2020
Available online 12 May 2020
1053-8119/© 2020 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
functional networks that share a common anatomical infrastructure (Friston, 1998).

Investigating functional networks that share anatomical areas requires models that explicitly allow for spatial overlap. One network model designed for the investigation of overlapping networks was proposed by (Smith et al., 2012), where distinct functional modes of brain activity were identified by means of their temporal independence. In their proposed model, time-series associated with different RSNs, as identified by spatial ICA (sICA), were further decomposed using temporal ICA (tICA) into a set of independent time courses, and associated node weights (i.e. RSN weights, also called the temporal ICA mixing matrix). For each independent time-course one can obtain an associated spatial map by linearly combining the original RSN maps using the node weights as coefficients of the combination. The authors refer to these maps as “temporal functional modes” (TFMs), instead of functional networks, after observing that TFMs contain significant amounts of anti-correlated activity - suggesting that modes may reflect the activity of more than a single network at a time. Smith et al. (2012) showed that TFMs were distinct from RSNs, and provide complementary information to that obtained with conventional FC analysis methods, making it possible to explore interactions between meaningful regions or networks of the brain (referred to as TFM nodes). Although tICA could be performed directly at the voxel level, such an approach would be both computationally inefficient and potentially less interpretable. Arguably, the most important contribution of TFMs is the mapping it provides between brain regions or networks, i.e., nodes, and a series of independent time-courses.

More recently, TFMs have been shown to be effective in identifying global structured noise in fMRI data that cannot be identified with spatially based decomposition techniques (Glasser et al., 2018). In particular, Glasser et al. (2018) suggested that the TFM model may be used for the removal of global physiological noise whilst retaining the global signal driven by neural activity, though this is still being debated (Power, 2019).

Despite these advances, the TFM analysis technique has hitherto not been widely used because it depends on tICA, which requires amounts of time-domain data not usually available in the majority of fMRI studies, which generally precludes the technique from being used at the single-subject level. In their studies, Smith et al. (2012) and Glasser et al. (2018) concatenated data from multiple subjects in order to compute the tICA decomposition. The reason why data were concatenated is because the ICA algorithm requires a large number of samples to function well. It requires many samples because it has to measure how much a signal deviates from normality, which is done by estimating higher-order moments of the signal distribution. According to a study by Herrmann and Theis (2007), and (see their Figs. 4 and 5, and Table 1), the recovery error of the ICA algorithm (FastICA) decreases with the inverse of the square root of the number of samples. An acceptable average recovery error of 0.1 can be achieved for most signal distributions with a minimal number of samples on the order of 3000. Because of this limitation, the TFM model has not been widely directly at the single-subject level.

Recently, with the development of new methods that push the temporal resolution of fMRI data such as Multiband or Simultaneous Multi-

Slice (Breuer et al., 2012; Feinberg et al., 2010; Larkman et al., 2001; Moeller et al., 2010; Setsompop et al., 2012), it is possible to obtain full-brain fMRI data with volume repetition times below 400 ms. Combining multiband with inverse imaging techniques, Hsu et al. (2017), obtained 20 image slices with a temporal resolution of 100 ms and a nominal spatial resolution of 5 mm isotropic by simultaneously exciting 10 imaging slices and using simultaneous echo refocusing (Feinberg et al., 2002). A recently developed technique (Boyacioglu et al., 2017) based on principles of inter-slice echo-shifting (Gibson et al., 2006) combined with SMS is capable of acquiring fMRI data with similar temporal resolutions but using conventional echo-planar imaging (EPI) reconstruction routines. These advances suggest that TFMs may well be feasible on data from a single fMRI session, albeit with limited spatial resolution.
A transition from a group analysis to single-subject measures is required to study inter-individual differences, for example mental health and behavioral measures (Smith et al., 2015). As such, the relevance of TFMs would increase significantly if the technique could be used at the single-subject level, making it possible to summarise high dimensional time-series into subject-dependent measures that could offer additional insights above those available from sICA. Here we explore this potential by using the MESH technique (Boyacioglu et al., 2017) and report TFMs obtained at the single-subject level from both task and resting-state fMRI.

2. Methods

2.1. Data acquisition

Six healthy volunteers (3M/3F, 32.0 ± 10.2 years old) participated in
the experiments after giving informed consent. The experiment was approved by the local ethics committee (CMO Arnhem-Nijmegen, the Netherlands). All data were acquired on a Siemens MAGNETOM Prisma (3T) MRI scanner (Erlangen, Germany) with a 32-channel head coil. Each participant was scanned in two 1-h sessions separated by a 15 min break.

Each session consisted of a resting-state run, a visual-motor association task, and a second resting-state run. Functional data were acquired with a Multiband Echo-Shifted (MESH) 2D EPI sequence (Boyacioglu et al., 2017) with the following parameters: TR = 158 ms, (effective) TE = 48 ms, resolution = 4.5 mm², 30 slices without gap, SMS = 6 with a CAIPI FOV/3 shift, no partial Fourier and no in-plane acceleration. The long TE achieved with echo shifting allowed us to accelerate even further by removing the fat saturation pulses (since the fat signal was highly attenuated at this long TE). Each resting-state scan lasted for 10 min and 45 s, corresponding to 4096 time-points, the maximum possible on the system, for a single run. In total, 43 min of resting-state data (16,384 timepoints), and 22 min of task-data (8192 timepoints) were collected from each subject. The flip angle was set to 7°, and in one subject to 6° to comply SAR limitations. Image reconstruction was performed online using the vendor-supplied implementation of Slice GRAPPA (Setsompop et al., 2012) with Leak Block (Cauley et al., 2014).

Anatomical data, for co-registration purposes, were acquired using a sagittal 1 mm isotropic MP-RAGE with a TR of 2300 ms, a TI of 900 ms, a TE of 3 ms, a flip angle of 9°, a turbo factor of 16 and an in-plane acceleration factor of 2 with a total acquisition time of 5:12 min. The anatomical scan was acquired at the end of the first session.

All imaging sequences were automatically aligned using a vendor supplied AutoAlign localizer sequence, which orients slices obliquely on a plane parallel to a line passing by the rostrum and splenium of the corpus callosum. During the scanning sessions participants lay supine in the MRI scanner, and with their eyes open. Subjects were awake throughout the whole duration of the experiment, as reported after questioning at the end of the scan. All datasets were collected between March and May 2018, and organized according to the Brain Imaging Data Structure (BIDS).

2.2. Visual motor association task

The visual motor association (VMA) task performed by the subjects consisted of learning arbitrary associations between 8 different Japanese Kanji characters (visual stimulus) and 4 different button presses (motor responses) (Grol et al., 2006). A visual stimulus was presented for 200 ms and subjects had to respond with a button press within 1.5 s after the stimulus onset. The motor response consisted of the flexion of one of four fingers of the right hand (a single left-handed subject participated in the experiment, yet was asked to use his right hand for the task) to press a button on a four-button fiber optic response pad - each button corresponding to two different Kanji characters. Subjects learned the associations between characters and buttons by trial and error. To know whether an association is correct, a visual feedback was given immediately after the response with a duration of 50 ms. The visual feedback displays either a green, red, or blue square, depending on whether the responses are correct, incorrect or exceed the 1.5 s reaction time (RT) cutoff, respectively. Trials were separated by an inter stimulus interval (ISI) taken from a random uniform distribution between 3.4 and 5.8 s, as exemplified in Fig. 1.

Presentation 20.1 (Neurobehavorial Systems, San Francisco, CA) was used to present the visual stimuli and register the motor responses. The visual stimuli were projected on a screen and viewed by the subjects via a mirror attached to the head coil. The subjects placed their index, middle, ring and little fingers on the corresponding button of the response pad, which were attached to the four corners of the subjects’ abdomen.

This particular task was chosen because it elicits reaction times on the order of 1 s (6 ± 2 TRs), and recruits both visual and motor areas, as well as attention and memory. In the current work these task datasets were analyzed to determine the effect of task on TFMs.

2.3. Data preprocessing

Data preprocessing was performed using FMRIprep version 1.0.15 (Esteban et al., 2018), a Nipype (Gorgolewski et al., 2011) based tool. Each T1w (T1-weighted) volume was corrected for INU (intensity non-uniformity) using N4BiasFieldCorrection v2.1.0 (Tustison et al., 2010) and skull-stripped using antsBrainExtraction.sh v2.1.0 (using the OASIS template). Spatial normalization to the ICBM 152 Nonlinear Asymmetrical template version 2009c (Fonov et al., 2009) was performed through nonlinear registration with the antsRegistration tool of ANTs v2.1.0 (Avants et al., 2008), using brain-extracted versions of both T1w volume and template.

Functional data were motion corrected using MCFLIRT (FSL v5.0 (Jenkinson et al., 2002)). “Field-map-less” distortion correction was performed by co-registering the functional image to the same-subject T1w image with intensity inversion (Huntenburg, 2014; Wang et al., 2017) constrained by an average field map template, implemented with antsRegistration (ANTS). This was followed by co-registration to the corresponding T1w image using boundary-based registration (Greve and Fischl, 2009) with 9 degrees of freedom, using FLIRT (FSL). Motion correcting transformations, field distortion correcting warp, BOLD-to-T1w transformation and T1w-to-template (MNI) warp were concatenated and applied in a single step using antsApplyTransforms (ANTS v2.1.0) using Lanczos interpolation.

Physiological noise regressors were extracted applying CompCor (Behzadi et al., 2007). Principal components were estimated for the two CompCor variants: temporal (tCompCor) and anatomical (aCompCor). A mask to exclude signal with cortical origin was obtained by eroding the brain mask, ensuring it only contained subcortical structures. Six tCompCor components were then calculated including only the top 5% variable voxels within that subcortical mask. For aCompCor, six components were calculated within the intersection of the subcortical mask and the union of CSF and WM masks calculated in T1w space, after their projection to the native space of each functional run. Frame-wise displacement (Power et al., 2014) was calculated for each functional run using the implementation of Nipype. In total, FMRIprep computes 32 confound regressors. These regressors were not used in the current experiment to denoise the functional time-series, but rather to measure whether they correlate with TFMs time-series. After registration and after computing confound regressors, datasets were high-pass filtered with a cutoff frequency of 1/100s using FSLMATHS (FSL). Many internal operations of FMRIprep use Nilearn (Abraham et al., 2014), principally within the BOLD-processing workflow (For more details of the pipeline see https://fmriprep.readthedocs.io/en/1.0.15/workflows.html). All datasets acquired for this experiment were visually inspected to guarantee satisfactory pre-processing.

Finally, each preprocessed run was decomposed into spatial independent components using FSL MELODIC. Although MELODIC offers automatic dimensionality estimates, it does so by comparing the eigenspectrum of the data against the eigenspectrum of a random matrix. The automatic dimensionality estimation algorithm works under the assumption that the noise in the data is not strongly temporally correlated (Beckmann and Smith, 2004), which is contravened owing to the short TR of the current study. When the assumption is not met, MELODIC overestimates the intrinsic dimensionality leading to an unusually large number of components, which causes component splitting. To partially overcome this problem we enforced a dimensionality of 120 after experimenting with ICA model orders of 70, 100, 120 and 150. The value of 120 was found to yield a reasonable compromise between finding meaningful components and unstructured noise. Since MELODIC orders components by variance explained, an adequate model order can be estimated by checking whether increasing the dimensionality introduces only new unstructured noise components. After hand classification of IC components into signal and noise following best practice guidelines described in the literature (Griffanti et al., 2017), datasets were non-aggressively denoised by removing unwanted components using

D.E.P. Gomez et al. NeuroImage 218 (2020) 116783
fal_regfilt, in alignment to what was performed in the original work of Smith et al. (2012). In our experience, the number of signal components identified on single runs ranged from 30 to 50, retaining around 30–40% of the variance explained by the ICs. Most noise components removed were identified as noise because of their high frequency content in the range of cardiac pulsation (or its harmonics) and their locations close to the base of the brain, which was indicative of cardiac physiological noise. All further processing and analysis is performed on the preprocessed and denoised datasets.

### 2.4. Temporal functional modes

In contrast to methods that identify patterns of activity based on spatial clustering approaches, such as co-activation pattern analysis, where patterns are identified by clustering similar fMRI time-frames using k-means (Liu and Duyn, 2013), or dynamic MVPA, where whole-brain images are clustered based on their spatial pattern correlations (Chen et al., 2018), the TFM model does not consider each frame to belong to a different spatial pattern or brain state, but rather to a superposition of latent independent states, each with its own time-course.

The TFM generation involves 2 steps: the first step consists of an initial spatial data reduction (or parcellation), where the fMRI data with V voxels and t timepoints is approximated by a product of d spatial maps (or parcels) and their associated time-courses, contained in the spatial mixing matrix sM, illustrated in equation (1) below. The second step consists of decomposing the spatial time-courses in sM, via temporal ICA, as a product of independent time-courses tIC and k node weights (2). The TFM model thus describes the data as the product of three matrices: the original parcels (or nodes), a mixing matrix consisting of node weights representing the link between space and time domains, and a set of independent time-courses (3). The product of the parcels and the node weights are used to generate the TFM spatial maps, such that the original data can also be seen as a set of TFM and their associated independent time-series (4).

\[
\text{Data}_{t,v} = \text{parcels}_{t,v} \times sM_{v,t} \\
sM_{v,t} \approx \text{weights}_{d,v} \times tIC_{d,t} \\
\text{Data}_{t,v} = \text{parcels}_{t,v} \times \text{weights}_{d,v} \times tIC_{d,t} \\
\text{Data}_{t,v} \approx \text{TFM maps}_{t,v} \times tIC_{d,t}
\]

In the original TFM publication, the initial maps and time-courses were obtained by means of a group sICA (gICA), such that nodes were chosen to be RSNs independent in space. Here we studied the dependence of the TFM model on the initial spatial data reduction step by comparing three different parcellation strategies:

1. A gICA decomposition, as done in Smith et al. (2012). To obtain subject-specific time-series we use dual-regression.
2. Individual sICA decompositions, and
3. An atlas-based decomposition using the Multi-resolution Intrinsic Segmentation Template (MIST) 444-parcel atlas (Urchs et al., 2017).

In the group sICA decomposition, the set of the single-subject resting-state denoised data were temporally concatenated and used as input for the group sICA with a dimensionality set to 120. From the 120 components, since input datasets have already been individually denoised. The spatial maps associated with these signals were hand-labeled and used as templates for a dual regression in order to obtain run- and subject specific signal time-courses (Seckmann et al., 2009; Filippini et al., 2009). First, for each functional run, each spatial map is regressed (as spatial regressors in a multiple regression) against the subject’s 4D space-time dataset. This resulted in a set of subject-specific timeseries, one per component. Next, those timeseries were variance normalised and regressed (as temporal regressors, again in a multiple regression) into the same 4D dataset, resulting in a set of subject-specific spatial maps. This second regression results in parameter estimate maps (regression coefficients), which are used as nodes for TFM. In the individual sICA approach, the sICA signal component time-courses were used directly as input to tICA. In the atlas approach, signals were obtained directly from the denoised datasets by averaging voxel time-courses within parcels of the 444-region brain atlas. We chose the MIST atlas because it offers a multi-resolution parcellation of the cortical, subcortical and cerebellar gray matter that provides annotated functional parcellations at nine resolutions from 7 to 444 functional parcels and provides an overlap-based pseudo-hierarchical decomposition tree that relates parcels across resolutions in a meaningful way. The pseudo-hierarchical labels can be used to describe TFM both in high resolution anatomical terms as well as low resolution functional terms.

For all 3 parcellation strategies, the demeaned, variance-normalised signal time-courses were then used as input to a tICA (FastICA, as implemented in scikit-learn (Pedregosa et al., 2011)) with a dimensionality set to 21. We report here a choice of model order equal to 21, as described previously in the literature (Smith et al., 2012).

#### 2.5. Data analysis

Our data analysis consisted of comparing TFM obtained with different parcellation strategies, measuring the correlation of TFM to confounding signals, and studying the within and between subject similarity of TFM as measured by correlations - for both the atlas as well as the group ICA parcellation approaches.

From the perspective of the temporal ICA algorithm, each parcel time-series is a mixture of unknown latent time-series. Different parcellation strategies are not constrained to have the same number of parcels, nor parcels of the same size, and thus may lead to vastly different input mixtures for ICA.

#### 2.6. Investigating the impact of parcellation choice

To verify whether the initial choice of parcellation influences the TFM spatial maps and independent time-courses, we compared TFM results obtained with gICA, with single-subject sICA and with atlas parcels. We compared TFM across these 3 parcellation strategies by computing the pairwise correlation between time-series, as we expected to find the similar signals regardless of the initial mixtures given to the temporal ICA algorithm.

#### 2.7. Accounting for potential spurious correlations

Because correlations are influenced by the number of samples in the data and their auto-correlation, a proper assessment of correlation values should therefore use a reasonable null model (Tong et al., 2019).

To assess spurious correlations between TFM obtained from different parcellations, we generated a null distribution as explained in the following example: Let gICA and atlas be the ith and jth TFM time-series obtained using the gICA and the atlas approach, respectively, from the first resting-state run from subject k. The null distribution consists of all correlation values between gICA and atlas, for all i, j excluding the correlation between “true” TFM pairs (the most similar TFM identified for a same run using different approaches). The distribution includes values obtained from all k subjects. By computing the cumulative distribution function from the null distribution we obtain a minimum correlation of 0.246 corresponding to a pseudo p-value of 0.05.

To assess spurious correlations between TFM obtained using the same parcellation, we generated a null distribution in the following manner: Let TFM and TFM be the ith and jth TFM time-series, respectively, obtained from the first resting-state run from subject k. The null distribution
is computed by generating all correlations between TFMi and a time-reversed TFMi, an operation which preserves the auto-correlation structure of the signal. By computing the cumulative distribution function from this null distribution we obtain a correlation of 0.124 and 0.142 for the atlas and gICA approaches, respectively, corresponding to pseudo p-values of 0.05.

2.8. Identifying common TFMs across subjects

To find the most common TFMs across functional runs we used a search and filter clustering approach. We first generated a list with correlation values between all pairs of TFMs (from all different runs). We then filtered this list to only include correlations above 0.591, the minimum value we found for which any two functional runs have at least one correlated TFM pair. This value is far larger than the threshold of significance described above. We then counted the number of times that a given TFM appears on the list. After finding the most common TFM, we excluded it and all of its pairings from the list, and searched for the next common TFM. TFMs whose time-courses correlated more than 0.4 with confound regressors were excluded from the analysis. The arbitrary threshold of 0.4 was determined after visual inspection, yet the choice is corroborated by the fact that the largest spurious correlation found in our null model was 0.396.

2.9. Low-dimensionality TFM visualisation

We then investigated whether it is possible to characterise subject heterogeneity looking at the correlations between node weights for the dual regression approach. To gain insight into the nature of TFMs and to characterise individual subjects and their heterogeneity, we used uniform manifold approximation and projection (UMAP) to visualise 2D embeddings of the input RSN time-series, the TFM time-series and the node weights. While common data reduction techniques (e.g. PCA) are capable of capturing the structure of the data, they can be difficult to interpret. With non-linear techniques such as UMAP we are capable of also capturing local structure (McInnes and Healy, 2018). We chose UMAP specifically because it generates embeddings that are more stable and consistent, and arguably of better quality than other non-linear dimensionality reduction techniques such as t-SNE (van der Maaten and Hinton, 2008) and FISHERMAP (Fischl et al., 1999) surface for illustration, with their associated independent time-courses shown as overlays. Although the time-courses obtained through the tfCA do not depend on the chosen initial dimensionality reduction, the associated TFM maps do serve to show the spatial extent of the TFM. Nevertheless, in all approaches we identify similar sets of anti-correlated brain regions and independent time-courses.

3. Results

3.1. On the choice of the initial parcellation

Fig. 2 shows the correlation between TFM time-series obtained with the group ICA and the atlas approach, on the left, and the time-series from the atlas approach for a single-subject on the right. From 21 TFM pairs, approximately half have a correlation coefficient of 0.5 or higher, and between 6 and 8 TFMs have a correlation greater than 0.75 (except for subject 5, where correlation values were smaller). Similar values were found when comparing to the sICA approach, albeit slightly lower, as the single-subject example on the right illustrates. Time-series with large peaks of activity that deviate from the mean and contain more low frequency content are apparently more easily identified, judged by the better agreement between different parcellations (see right panel in the figure).

Fig. 3 illustrates 3 examples of highly correlated TFMs obtained with the parcellation strategies considered (time-course correlation > 0.9). For visualisation, TFM spatial maps in MNI space were projected into the Freesurfer fsaverage5 (Fischl et al., 1999) surface for illustration, with their associated independent time-courses shown as overlays. Although the time-courses obtained through the tfCA do not depend on the chosen initial dimensionality reduction, the associated TFM maps do serve to show the spatial extent of the TFM. Nevertheless, in all approaches we identify similar sets of anti-correlated brain regions and independent time-courses.

3.2. Correlation to confounds

We again used correlation coefficients to investigate how much TFM time-series are impacted by the 32 confound regressors such as e.g., the global signal, subject motion, and cardiac and respiratory fluctuations.

To illustrate TFM time-courses and their correlation to confound regressors, Fig. 4 shows data from an exemplary single subject dataset. The figure shows that confound contamination is widespread over time-series of atlas parcels, but can be parsimoniously extracted into a small number of independent TFMs, such as one widespread over the whole brain and highly correlated to the global signal. We should note that TFMs maps that reflect the global signal are marked by widespread activation throughout the brain, similar to the global TFMs described in Smith et al. (2012) and Glasser et al. (2018). For ultra-fast fMRI data, it is also possible (for some subjects) to identify cardiac TFMs with activation localized to the base of the brain, even after spatial ICA denoising.

3.3. TFM repeatability

Fig. 5 shows a graph representing the correlation structure of TFMs both within and between subjects for the gICA and the MIST atlas parcellation. In each graph, each of the 36 graph vertices correspond to one of the 36 functional datasets acquired in the experiment, and subjects are color coded as indicated in the figure. The edges connecting vertices are weighted by the count of pairs of TFMs with a correlation greater than 0.687 (for the gICA) or 0.591 (for the atlas, as used in the previous section), i.e., the minimum value for which the graph is complete (i.e. 0.687 for the gICA and 0.591 for the atlas). As described previously, the correlation between TFMs is measured as the correlation between the rows of the temporal ICA mixing matrix i.e. the node weights. Because TFMs are unordered, each row of each dataset is compared to all rows of other datasets such that there are in total $21 \times 21 = 441$ comparisons per pair of datasets.

We find that, similar to resting-state networks, TFMs are subject specific and repeatable across sessions. This is highlighted by heavier edges within subject clusters in the force-directed graph.

In fact, even without any thresholding, the average correlation within subjects is found to be larger than between subjects, as shown in Table 1.
This observation holds true for all subjects. Here, on average, the ratio of within to between subject correlations between TFMs was 1.86 for the dual regression approach and 1.21 for the atlas approach.

### 3.4. Most commonly identified TFMs

To compare TFMs obtained with the same parcellation in different subjects and sessions, we computed the correlation between node weights, as this is both inexpensive computationally and more robust than measuring for example Dice coefficients between TFM maps. We choose to display common TFMs obtained with the atlas approach for three reasons. First, the anatomical atlas parcels are non-overlapping. A voxel either fully belongs to a parcel or it does not. This is in contrast to the gICA approach, where different soft parcels (in contrast to hard parcels, where voxels either do or do not belong to a parcel) may slightly intersect and voxels have different weights in each parcel. Second, the gICA template was generated with data from our participants, and so it is exclusive to this study. Third, our gICA template labels may not correspond perfectly to the canonical networks described in the literature. The MIST atlas, on the other hand, is not exclusive to this study and the labels are well defined. The MIST atlas pseudo-hierarchical parcel information also makes it possibly to simplify the description of common TFMs by summarizing maps in terms of interactions between 36 higher-level functional parcels, instead of the higher resolution 444 parcel with which TFMs are computed.

With our approach to find similar TFMs across datasets we identified 6 modes that are common across the majority of runs, shown in Fig. 6. The most common mode was found in 31 out of 36 runs, and the sixth most common was found in 9 of 36 runs (with a correlation larger than 0.591). For simplicity we describe the node weights of commonly found TFMs by mentioning whether the DMN is present and then describing regions correlated or anti-correlated to it.

The most common TFM (Panel A) consisted of an anti-correlation pattern between the DMN and task-positive networks. The principal regions belonging to this TFM include the lateral DMN in anti-correlation to the postcentral sulcus and supra-marginal gyrus, the auditory network, posterior and anterior insula, visual networks and visual streams, and motor cortices. We identify this mode thus as being a default temporal mode.

The second most common TFM found amongst all participants (Panel B) shows a strong recruitment of the DMN, task-control and executive control network in anti-correlation to the auditory network and posterior insula, the somatomotor network, visual networks, amygdala and hippocampus. This TFM is suggestive of sensory integration, given that it most areas of the sensory system and limbic system are recruited simultaneously. Its timeseries shows bursts of activity that last for approximately 5–10s, or about 30 time-points. The third TFM (Panel C) displays the DMN, insula, thalamus and caudate nucleus, executive control network and ventrolateral prefrontal cortex in opposition to all areas of the visual network and visual streams. Because of the pronounced presence of visual areas, we identify this TFM as being a visual temporal mode. The fourth TFM (Panel D) displays almost no presence of the DMN. It shows instead a recruitment of the somatosensory network, premotor and supplementary motor cortices, the superior posterior cerebellum (hemispheric lobule VI), the fronto-parietal task control network and the executive network. This “motor” TFM displays slight anti-correlation to the visual areas and middle temporal gyri. It also displays, in contrast to other TFMs, signs of hemispheric asymmetry. We identify this TFM as being a motor temporal mode. We note that, when we analyze task and resting-state datasets separately (see supplementary materials), this TFM is the most commonly found in the task datasets. This is likely because our visual-motor association task recruits the motor and visual areas.

The fifth TFM (Panel E) displays little presence of the DMN. This TFM recruits lateral areas of the DMN, the visual networks and visual stream, the amygdala and hippocampus in opposition to the auditory network and insula, the ventrolateral somatomotor network, task control and executive networks. Finally, the sixth TFM (Panel F) integrates areas of the fronto-parietal task control network, ventrolateral prefrontal cortex, dorsal and ventral visual streams, the lateral DMN, and the executive networks.
control network in strong opposition to visual and somatomotor networks, as well as anterior cingulate and ventromedial prefrontal cortices. Because data were pooled from both rest and task datasets, we present analysis results separated by task and rest in the supplementary materials. There we find that the lateraliased motor control network is most commonly found in the task datasets.

3.5. Characterising subjects according to node weights

To characterise individual heterogeneity, we looked at UMAP embeddings of the original signal timeseries and their TFM decomposition. The embeddings illustrate how TFMs are capable of highlighting between-subject differences in a manner that is qualitatively different from what is possible with spatial ICA alone, and are shown in Fig. 7. In the figure, each of the three scatter plots represent UMAP embeddings generated from 1. dual-regression (gICA) RSN time-series, 2. TFM node weights and 3. the TFM time-series, for all runs and subjects.

Although the axis of the 2D plots are not directly interpretable (any rotation would be an equally valid embedding), dots that are similar, as measured by correlation, in the original space, are adjacent in the 2D space. Should there be any underlying subject bias in the RSN time-courses, where within-subject time-courses are more similar than between-subject time-courses because of subject-specific idiosyncrasies, for example, then UMAP should bring this out in the form of clusters in the plot. With this in mind, it is possible to see that the RSN time-series embedding (left plot) does contain subject-specific clusters (dual-regression permits the identification of between-subject differences (Bockmann et al., 2009)) yet not to an extent where studying subject heterogeneity would be readily possible - in the embedding only a subset of clusters can be readily disentangled. With the TFM analysis, the subject structure that is imprinted in the RSN time-series can be clearly extracted into the node weights of the TFM model (middle plot), whereas no subject-specific content is present in the independent time-series (right plot). In the node weights plot (middle), each dot corresponds to the node weights vector of a single TFM. The plot shows that the node weights matrix, which contains the link between space and the independent time-courses, is highly subject specific.

3.6. Task perturbations on TFMs

Finally, Fig. 8 illustrates TFM differences between subjects and between resting-state and the VMA task. In the figure, two plots are shown side-by-side. In both plots dots in the scatter plot represent node weights, just as in 7. On the left, node weights are color-coded by subject number, whereas on the right they are coded by task: rest-1 is the first resting-state obtained in a session, which is following by the task (VMA task), and rest-2, a second resting-state acquisition. The figure suggests that task activity induces but a small perturbation in brain network organisation, and that these effects are dwarfed by subject specific differences.

Fig. 7. Left: Each dot represents a signal time-series (59 per run, 6 runs per subject). Middle: Each dot represents an RSN to TFM mapping, i.e., a row of the tICA mixing matrix (21 per run, 6 runs per subject). Right: Each dot represents an independent time-series (21 per run, 6 runs per subject). In all plots, dots are color coded by subjects. There is no interpretability to the y and x axis, as any rotation would represent an equally valid embedding.

Fig. 8. UMAP applied to all RSNs to TFMs mappings, i.e., rows of the tICA core mixing matrix. Each dot represents one RSN to TFM mapping, yielding 21 (TFMs) * 3 (runs) dots per subject. Both plots contain the same underlying information, yet color coded by subjects (on the left) or task (on the right). Tasks rest-1, task and rest-2 correspond to the three functional runs that are acquired in order in each scanning session. The axis are arbitrary, as any rotation would represent an equally valid embedding.

Direct correlations between TFM time-courses and task regressors indicate that TFMs do carry an imprint of the task design, as shown in Fig. 9. The distribution of these correlations suggest that the task design is imprinted in multiple TFMs, rather than eliciting the appearance of specialized task TFMs. This is seen by observing a general flattening of the distribution. Nonetheless, TFMs with a high correlation to task regressors do present a strong loading in motor areas, as expected from the task design. This is shown in Fig. 10, which compares a single-subject thresholded TFM map and a GLM activation map. The maximal correlation found between a TFM and the task regressor, for each subject and session, is reported in Table 2.

4. Discussion

In the current contribution we demonstrated the feasibility of extracting temporally independent functional modes at the single-subject level by leveraging an ultra-fast fMRI sequence capable of sampling the whole-brain with isotropic resolution every 158 ms. Our findings suggest that the technique is informative at the single-subject level.
that TFMs can now be obtained at the single-subject level, we believe that temporal ICA may expand the reach of ICA based denoising techniques further than is possible with spatial ICA alone. As an example, cardiac signal was still identified with temporal ICA even after regressing out multiple cardiac components during the spatial ICA denoising. A drawback of temporal ICA for denoising, though, is that it requires the artifactual sources to be independent of the signals of interest. In the case of a motion artifact that is time-locked to a stimulus, for example, the TFM model would not be optimal for the denoising procedure.

4.3. Repeatability

We investigated whether TFMs are stable by applying the model to datasets from different runs and sessions and measuring correlations between node weight vectors across datasets. Our analysis showed that TFMs are repeatable and subject specific, suggesting that the method is robust at the single-subject level. This effect is independent of the initial spatial dimensionality reduction, and was present in both the gICA and the atlas approach, albeit more pronounced in the group ICA approach, likely because gICA provides a basis derived directly from the cohort of participants of this particular study. Unfortunately, we could not directly test the repeatability of the individual spatial ICA approach because the number of spatial signals and their ordering differ between datasets, which precludes the computation of correlation estimates between node weights.

4.4. Commonly identified TFMs

Although our functional runs had a short duration of 10:45 min, we could detect six common TFMs that were present in multiple runs from distinct subjects. The detection of TFMs in a 10-min run is limited not only by the number of samples acquired, but also by the fluctuations that can be observed in such a short period of time. In traditional functional connectivity studies, the longer a subject is sampled, the more stable and identifiable are functional networks (Laumann et al., 2015). Although group averaging allows reproducible identification of a larger number of temporal functional modes (Glasser et al., 2018) than described in the current contribution, averaging may not be feasible for studies with a small number of subjects, or studies where the time allocated for functional imaging is limited. Nonetheless, despite runs of short duration, we could identify biologically plausible TFMs that reflect commonly observed patterns of activity, such as the anti-correlation pattern between the DMN and the task-positive networks, the integration of the visual network and visual streams, as well as an integration of motor areas including the cerebellum. Other temporal functional modes identified may hint at modes of sensory integration that are not generally observed with correlation analysis or spatial ICA. This is highlighted by looking at the peaks in time-series of the putative sensory integration TFM, and the motor TFM. This feature is seen in TFMs that recruit mostly sensory areas with little to no presence of the DMN. The peaks last for approximately a BOLD cycle of 5–10 s, or 30–60 frames, which is too long a time to correspond to MR system imperfections or motion. The duration between peaks ranges from 100 to 300 s. They could represent interferences from BOLD oscillations with different frequencies on the order of 0.1 Hz. Finally, we note that even within the common TFMs there is between-subject variability, so although we interpret these areas as observations of possible “canonical” TFMs, further study with different datasets and populations, and different acquisition methods would be necessary to corroborate our findings. Additionally, not all 6 common TFMs are observed in every subject, similar to how not every RSN is found in every single subject in a spatial ICA decomposition.

4.5. Specificity of node weights

We showed that it is possible to characterise subject heterogeneity by investigating correlations between node weights obtained with the group

---

**Table 2**

| Subject | Session | Correlation to task regressor |
|---------|---------|-------------------------------|
| 1       | 1       | 0.17                          |
| 2       | 1       | 0.30                          |
| 3       | 1       | 0.25                          |
| 4       | 1       | 0.23                          |
| 5       | 1       | 0.31                          |
| 6       | 1       | 0.28                          |

**4.1. Choosing a parcellation**

We hypothesized that, provided the granularity of the initial decomposition is finer than the resolution of the underlying TFMs to be studied, any meaningful parcellation should be acceptable. We showed that, indeed, reproducible modes of activity can be extracted irrespective of the choice of initial parcellation. This result is in agreement with our expectations, as the ICA algorithm should identify latent signals independent of their coefficients in the original mixed signals. The number and size of the initial parcels, however, does influence the smoothness and boundaries of the resulting TFM maps, and also limits the number of TFMs that can be extracted. Nonetheless, the fact that there is an equivalence in the time-series obtained from different data reduction steps, as shown in Figs. 2 and 3 shows that we are free to choose any meaningful basis set to describe temporal functional modes. By looking at TFMs as mixtures of RSNs obtained from a group ICA (and dual regression) on resting-state data, for example, we can study the heterogeneity of subjects within a group. The TFM core mixing matrix gathers all subject specific information present in the original RSN signals, which may be driven by idiosyncrasies of subject’s anatomy. By looking at TFMs as mixtures of anatomical regions of the brain we can simplify the description of the resulting TFM maps, i.e., it is easier to reason about combinations of parcels without any overlap, such as parcels from an atlas, and identify anatomically similar TFMs across subjects. This is because with a hard parcellation each brain region is represented by at most one node weight, whereas with a soft parcellation a brain region may be represented by multiple node weights according to the amount of node overlap.

**4.2. Correlation to confounds**

As shown in Glasser et al. (2018), temporal ICA is capable of clearly separating artifactual sources from BOLD fluctuations of interest. Given

---

Fig. 10. A comparison between an activation map (Gaussianized z-scores) obtained with a General Linear Model and the overall TFM with highest correlation (0.44) to a task regressor (TFM 14, subject 5, session 2). Both maps are seen to recruit sensory-motor as well as supplementary motor areas, and show lateralization. The TFM map, however, shows anti-correlated areas of the default mode network.
ICA and dual regression approach. We were able to visualise subject-specific fingerprints by using a non-linear dimensionality reduction technique that attempts to preserve both local and global structure, UMAP. We showed that the between-subject differences initially present in the RSN time-series can be imprinted in the node weights in the TFM model. These results show that node weights can be used as powerful descriptors for the characterization of subject heterogeneity. Caution must be exercised, however, to not over interpret UMAP results. In contrast to linear dimensionality reduction techniques such as PCA, the axis of UMAP embeddings have no meaning. The tuning of hyper-parameters may increase or decrease the distance between points in the embeddings, and, furthermore, distances cannot be compared across different embeddings. Finally UMAP may find artificial structure even in noisy datasets (McInnes and Healy, 2018).

4.6. Task perturbations

When studying the influence of task against the resting-state, we found evidence that external stimuli produced only a minor perturbation on all TFMs, instead of eliciting exclusively task-related TFMs. This is supported by the small and general increase in correlation values between TFMs and task regressors (Fig. 9), and by negligible differences between task and rest datasets in the node weight embedding (Fig. 8, left panel). Differences between node weights computed from rest and task datasets were found to be much smaller than the difference between subjects, a result similar to what had been observed in the context of resting-state networks (Gratton et al., 2018). If true, these observations support the idea that task-evoked activity is associated with general shifts in overall network organisation, rather than changes in functional connectivity between specific brain regions alone (Bolt et al., 2017). Additionally, our findings corroborates a recent report from Taglia et al. (2018), who suggest that latent brain states may not necessarily be aligned with task conditions, and call for unsupervised algorithms for estimating dynamic states and the transitions between them. Temporal ICA, being an unsupervised algorithm, may be a helpful tool to explore these states. Nonetheless, as pointed out by Calhoun et al. (2001), the characteristics of the latent signals, in particular in what concerns their temporal and spatial distribution, will define whether they can be unambiguously identified by spatial and temporal ICA. The high correlation found in one run between TFM and task regressors, also suggests that if TFMs load highly on areas driven by task activity, and if the task activity is temporally independent from other underlying cognitive processes, then some TFMs may indeed resemble task activation maps (Fig. 10). Because of the robustness of the TFM algorithm in identifying independent time-courses regardless of initial parameters, we expect it to be an additional useful tool for the modeling of dynamic connectivity. The study of interactions between networks, which may follow from investigating the behaviour of TFM node weights in rest and task states, could potentially be relevant both for clinical as well as neuro-scientific applications, even at the single-subject level.

4.7. Limitations

We acknowledge that there are limitations to our work. First, the proposed ultra-fast sequence requires compromises in image quality and spatial resolution, and therefore may not be adequate for investigations of small structures of the brain, as those would benefit from a higher spatial resolution. Additionally, the low spatial resolution cannot simply be overcome by imaging at higher field-strengths with conventional gradients, as the long echo-time associated with the MESH sequence would be prohibitive. Second, there are limitations inherent to the ICA algorithm with respect to temporally auto-correlated signals, which may increase the recovery error (Herrmann and Theis, 2007). There must be, thus, a limit where faster temporal sampling starts to yield diminishing returns, but this limit - in part driven by the speed with which oscillations in the haemodynamic response can be observed - is yet to be determined.

4.8. Future research

We plan to use the TFM model to investigate group differences between healthy controls and patients with psychiatric disorders, in a similar fashion as we have done in the current work for the investigation of subject heterogeneity. Additionally, further advances in sequence development and image reconstruction may allow TFMs to be used for the investigation of latent signals in finer parcellation including subcortical areas. Finally, studies with a larger number of participants would be required to investigate whether there are more temporal functional modes common across populations.

Acknowledgements

Daniel Gomez is funded by a Marie Curie FP7-PEOPLE-2013-ITN “Initial Training Networks” Action from the European Union, and the grant ID is FP7-PEOPLE-2013-ITN (Project Reference Number: 608123). We are thankful to Zahra Fazal and Samantha Noteboom for help during data collection and preparation of the task paradigm. We are thankful to Thomas Beck and Benedikt Poser for support during sequence development. We are thankful to Steven Smith (Oxford University) for comments and suggestions during the early stages of this project.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.neuroimage.2020.116783.

References

Abraham, A., Pedregosa, F., Eickenberg, M., Gervais, P., Muller, A., Kossaifi, J., Varoquaux, G., 2014. Machine learning for neuroimaging with scikit-learn, 8 (1-10) pmid: 24600388.
Avants, B., Epstein, C., Grossman, M., Gee, J., 2008. February. Symmetric diffeomorphic image registration with cross correlation: evaluating automated labeling of elderly and neurodegenerative brain. Med. Image Anal. 12 (1), 26-41.
Beckmann, C.F., Smith, S.M., 2004. Probabilistic independent component analysis for functional magnetic resonance imaging. IEEE Trans. Med. Imag. 23 (1), 137–152. pmid: 14964560.
Beckmann, Mackay, Filippini, Smith, 2009. Group comparison of resting-state fMRI data using multi-subject ICA and dual regression. Neuroimage 47, S148 pmid: 19357304.
Behzadi, Y., Restom, K., Liu, J., Liu, T.T., 2007, August 1. A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. Neuroimage 37 (1), 90–101.
Bolt, T., Nomi, J.S., Rubinov, M., Uddin, L.Q., 2017, April. Correspondence between evoked and intrinsic functional brain network configurations: functional Brain Network Configurations. Hum. Brain Mapp. 38 (4), 1992–2007.
Boyacioglu, R., Schulz, J., Norris, D.G., 2017. Multiband echo-shifted echo planar imaging. Magn. Reson. Med. 77 (5), 1981–1986 pmid: 27297682.
Breuer, F., Blaimer, M., Griswold, M., Jakob, P., 2012. Controlled aliasing in parallel imaging results in higher acceleration (CAIPIRINHA). MAGNETOM Flash 135–142.
Calhoun, V., Adali, T., Pearlson, G., Pekar, J., 2001. Spatial and temporal independent component analysis of functional MRI data containing a pair of task-related waveforms. Hum. Brain Mapp. 13 (1), 43-53.
Cauley, S.F., Polimeni, J.R., Bhar, H., Wald, L.L., Setsompop, K., 2014. Interlase leakage artifact reduction technique for simultaneous multislice acquisitions. Magn. Reson. Med. 72 (1), 93–102 pmid: 23963964.
Chen, R.H., Itu, T., Kulkarni, K.R., Cole, M.W., 2018, September. The human brain traverses a common activation-pattern state space across task and rest. Brain Connect. 8 (7), 429–443.
