



Total activation: fMRI deconvolution through spatio-temporal regularization

Fikret Işık Karahanoğlu^{a,b,*}, César Caballero-Gaudes^{b,c}, François Lazeyras^b, Dimitri Van De Ville^{a,b}

^a Institute of Bioengineering, Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland

^b Department of Radiology and Medical Informatics, University of Geneva, Switzerland

^c Basque Center on Cognition, Brain and Language, Donostia-San Sebastian, Spain

ARTICLE INFO

Article history:

Accepted 22 January 2013

Available online 4 February 2013

Keywords:

fMRI
BOLD
Deconvolution
Sparsity
Total variation
Spatio-temporal regularization
Paradigm-free mapping
Resting state

ABSTRACT

Confirmatory approaches to fMRI data analysis look for evidence for the presence of pre-defined regressors modeling contributions to the voxel time series, including the BOLD response following neuronal activation. As more complicated questions arise about brain function, such as spontaneous and resting-state activity, new methodologies are required. We propose total activation (TA) as a novel fMRI data analysis method to explore the underlying activity-inducing signal of the BOLD signal without any timing information that is based on sparse spatio-temporal priors and characterization of the hemodynamic system. Within a variational framework, we formulate a convex cost function—including spatial and temporal regularization terms—that is solved by fast iterative shrinkage algorithms. The temporal regularization expresses that the activity-inducing signal is block-type without restrictions on the timing nor duration. The spatial regularization favors coherent activation patterns in anatomically-defined brain regions.

TA is evaluated using a software phantom and an event-related fMRI experiment with prolonged resting state periods disturbed by visual stimuli. The results illustrate that both block-type and spike-type activities can be recovered successfully without prior knowledge of the experimental paradigm. Further processing using hierarchical clustering shows that the activity-inducing signals revealed by TA contain information about meaningful task-related and resting-state networks, demonstrating good abilities for the study of non-stationary dynamics of brain activity.

© 2013 Elsevier Inc. All rights reserved.

Introduction

Conventional analysis of functional magnetic resonance imaging (fMRI) data heavily relies on approaches based on general linear models (GLMs) where prior knowledge about the experimental paradigm; i.e., onsets and durations of stimuli, is used to construct temporal regressors, which are then fitted to the time course of every voxel. The subsequent statistical hypothesis testing, for a given contrast of fitted weights, is a mass univariate approach leading to an “activation map” that highlights brain regions for which sufficient evidence is present to be related to the experimental paradigm. The relationship between the measured blood oxygen level dependent (BOLD) and the experimental paradigm can be modeled as a linear shift-invariant system with the hemodynamic response function (HRF) as an impulse response. In standard GLM approaches (Friston et al., 1998), the HRF is predefined using gamma functions; more flexible techniques estimate HRF components in a subject- or time-dependent way in order to deal with inter- and intra-subject variability (Aguirre et al., 1998). More notably, parcel-

based HRF estimation methods through joint detection estimation (JDE) framework are studied based on Bayesian approaches (Chaari et al., 2013; Makni et al., 2008; Vincent et al., 2010). Recently, an adaptive parcel identification driven by the hemodynamics is proposed using JDE (Chaari et al., 2012; Thirion et al., 2006). These HRF identification methods are mainly combined with GLM analysis to explore the parcel/subject/group/task specific hemodynamic models.

Not all brain activity can be modeled beforehand using stimulus functions; e.g., interictal epileptic discharges occur spontaneously. In addition, mere resting-state, which was neglected before as background noise, is known (Raichle, 2006) to produce characteristic patterns of brain activity referred to as resting-state networks. Such unpredictable activity cannot be inferred from traditional GLM analysis approaches (Gusnard and Raichle, 2001). Therefore, there is an increasing need for methodologies that enable the exploration of hemodynamic brain activity without predefined responses (Cole et al., 2010). Data-driven methods have been proposed for that purpose such as fuzzy clustering (Baumgartner et al., 2000), temporal clustering analysis (TCA) (Liu et al., 2000; Morgan et al., 2008), seed correlation analysis (Biswal et al., 1995), or subspace decomposition methods such as independent component analysis (ICA) (Beckmann and Smith, 2004; Calhoun and Adali, 2006), canonical correlation analysis (CCA) (Afshin-Pour et al., 2012) and agnostic canonical variates analysis (agnostic-CVA) (Evans et al., 2010). ICA is probably the

* Corresponding author at: Institute of Bioengineering, Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland.

E-mail address: isik.karahanoglu@epfl.ch (F.I. Karahanoğlu).

most commonly used data-driven method. It provides a bilinear decomposition of the data into components that consist out of a spatial map with an associated time course. Its application to fMRI typically relies on (a surrogate for) spatial statistical independence between the components. However, such criterion is not directed specifically to identify “activation like” components since no knowledge is taken into account about the hemodynamics or about the type of activity-driven signal (e.g., spikes versus sustained activity).

fMRI deconvolution methods have been proposed to uncover the underlying activity-inducing signal at the fMRI timescale of seconds. Initially, Glover (1999) introduced Wiener deconvolution filtering that is optimal for Gaussian sources and thus results in very smooth activity-inducing signals. This work was generalized by Gitelman et al. (2003) to study the psychophysiological interactions at the neuronal level. Following recent advances in convex optimization theory, these methods can be made more sophisticated by adding sparse priors on the underlying signal and solve maximum a posteriori estimation rather than naive maximum likelihood estimation. They have been exploited as an extension to standard GLM analysis by defining spatial priors on spatial activation maps (Flandin and Penny, 2007; Harrison et al., 2008; Smith and Fahrmeir, 2007; Vincent et al., 2010). Within a temporal fMRI deconvolution framework, while the linear system assumption on the hemodynamic model is retained, regularization terms use ℓ_1 -norm to favor sparse solutions in time; i.e., a limited number of spike-like activations. For example, Hernandez-Garcia and Ulfarsson (2011) use the majorization–minimization scheme, Gaudes et al. (2011, 2013) rely on ridge-regression and sparsity-promoting estimators, and (Khalidov et al., 2007, 2011) impose sparsity in the activelet-domain, a wavelet basis that is tailored to the hemodynamic properties. These methods do not require any knowledge on the timing and exploit temporal properties of the HRF. Promising results have been demonstrated for local event detection (Gaudes et al., 2011), especially in epilepsy (Lopes et al., 2012), but also for resting-state analysis (Petridou et al., 2012). Recently, non-linear models have also attracted a lot of attention for blind deconvolution to explore the network dynamics. These methods solve non-linear state-space model in continuous time, which bring along high resolution solutions, and access both hidden states and (non)dynamic parameters related to the neuronal activity via Bayesian filtering, Cubature Kalman filtering and Local Linearization filters (Friston et al., 2008, 2010; Havlicek et al., 2011; Riera et al., 2004). These methods, however, have high computational cost compared with linear models and are mostly applied for uncovering the hemodynamics of a priori regions of interest.

Here we propose a novel deconvolution method for fMRI data analysis, for which we coin the term “total activation” (TA). TA includes some unique features to go beyond the limitations mentioned above:

1. *Express temporal properties of the activity-inducing signal.* The deconvolution identifies the “innovation” signal (which is spike-type) as the sparse driver of the BOLD signal. However, the activity-inducing signal can be more flexible such as block-type signals (Karahanoğlu et al., 2011).
2. *Structured sparsity for combined temporal and spatial regularization.* Spatial regularization is incorporated using mixed-norms based on anatomical priors of brain regions (Baritoux et al., 2011); i.e., time courses of voxels in the same brain regions are favored to be coherent.
3. *Take advantage of efficient optimization schemes.* We employ the efficient generalized forward–backward splitting algorithm (Raguet et al., 2012), which is a fast iterative shrinkage algorithm that alternates between temporal and spatial domain solutions until convergence to the final estimate of the underlying activity-inducing signal.

The paper is organized as follows. We first introduce the TA theoretical framework. Next, the feasibility of TA is demonstrated on both synthetic and experimental data. The simulation study allows validating the performance for block-type activity-inducing signals with different block durations. The experimental study is based on fMRI data acquired in three healthy subjects while being at rest but with several (unexpected) visual stimuli. The TA deconvolved signals indicate strong and short periods of activity in the primary visual regions that match with the experimental timing. The dynamic “activation maps” show coordinated activation (and “de-activation”) in large-scale networks, mostly with much larger average block lengths. We also show that resting-state networks can be retrieved using hierarchical clustering of the region-averaged activity-inducing signals.

Total activation

fMRI signal model

We represent the BOLD response following neuronal activation as the convolution of the activity-inducing signal $u(t)$ with the HRF $h(t)$. In model-based approaches, the activity-inducing signal corresponds to the stimulus function according to the experimental paradigm. Hence, for every voxel i it can be modeled by a weighted sum of shifted and dilated box functions $b(t)$ as

$$u(i, t) = \sum_k c_k(i) b(t/a_k - t_k), \quad (1)$$

where $b(t) = 1$, $0 \leq t \leq 1$ and 0 otherwise; $c_k(i)$ is the amplitude of the k -th block; a_k is the block length; $a_k t_k$ is the onset timing of activity. We then define the innovation signal $u_s(i, t)$ as the derivative of the activity-inducing signal. In particular, for block-type signals, the innovation signal will be sparse and contain many zeros as

$$\begin{aligned} D\{u(i, \cdot)\}(t) &= \sum_k c'_k(i) (\delta(t - a_k t_k) - \delta(t - a_k(t_k + 1))), \\ &= \sum_k c'_k(i) \delta(t - t'_k) = u_s(i, t), \end{aligned} \quad (2)$$

where D is the derivative operator and $\delta(t)$ is the Dirac-delta function. We have reparameterized the innovation signal to clearly reflect its sparse nature; i.e., a train of Dirac impulses. Hence, $u_s(i, t)$ represents the timing when the activity inducing signal $u(i, t)$ changes its amplitude. We further assume the following linear-system relationship between the activity-inducing signal $u(i, t)$ and the activity-related signal $x(i, t)$:

$$x(i, t) = u(i, t) * h(t), \quad (3)$$

where $h(t)$ is the hemodynamic response function (HRF). The canonical HRF used in SPM (Friston et al., 1998) is characterized by two gamma functions. An alternative formulation is the first-order Volterra series approximation of the Balloon model for fMRI BOLD (Friston et al., 2000; Khalidov et al., 2011). Here we use the linear differential operator L_h defined in Khalidov et al. (2011), which inverts the hemodynamic system; i.e., we have

$$L_h\{h\}(t) = \delta(t). \quad (4)$$

Then, we recover the activity-inducing signal as

$$L_h\{x(i, \cdot)\}(t) = u(i, t), \quad (5)$$

and, given the link between innovation and activity-inducing signal, we also have $L\{x(i, \cdot)\}(t) = D\{u(i, \cdot)\}(t) = u_s(i, t)$, where the operator $L = DL_h$ combines L_h with the regular derivative. More specifically,

the differential operator L_h is defined by its zeros α_i ($i = 1, \dots, M_1$) and poles γ_j ($j = 1, \dots, M_2$) as follows

$$L_h = \prod_{i=1}^{M_1} (D - \alpha_i I) \left(\prod_{j=1}^{M_2} (D - \gamma_j I) \right)^{-1},$$

where I is the identity operator and $M_1 > M_2$. In Fig. 1, we illustrate our fMRI signal model and its underlying sparse structure.

In practice, the activity-related signal $x(i, t)$ is corrupted by different noise and artifactual sources, such as non-neurophysiological contributions (e.g., aliased cardiac and respiratory fluctuations), movement, scanner drifts and thermal noise (Lund et al., 2006). The fMRI signal model $y(i, t)$ then becomes

$$y(i, t) = u(i, t) * h(t) + \sum_k \beta_k(i) n_k(t) + \epsilon(i, t), \quad (6)$$

where $n_k(t)$ represent known nuisance regressors (e.g., movement, low-frequency drifts), β_k are associated weights, and $\epsilon(i, t)$ is independently distributed Gaussian noise with zero mean and variance σ_ϵ^2 .

We represent the sampled and discretized full dataset as a matrix $\mathbf{y} = [y(i, t)]_{i,t}$ of size $N \times V$, where N is the number of scans and V is the total number of voxels. Also the operators need to be discretized, which we denote with Δ ; e.g., Δ_D indicates the finite-difference for the derivative D , and Δ_{L_h} for the hemodynamic inverse filter (Karahanoglu et al., 2011).

Variational formulation

TA aims at reconstructing activity-related signals from noisy fMRI measurements by imposing informative priors on the signal of interest using a variational formulation (Kirsch, 1996; Zibulevsky and Elad, 2010). Within the context of fMRI data processing, we introduce a novel spatio-temporal formulation based on the minimization of a cost function that includes a least-squares data-fitting term equal to the residual sum of squares (RSS), and two regularization terms \mathcal{R}_T and \mathcal{R}_S that act along the temporal and spatial dimensions, respectively. Specifically, our cost function reads

$$\tilde{\mathbf{x}} = \arg \min_{\mathbf{x}} \frac{1}{2} \|\mathbf{y} - \mathbf{x}\|_F^2 + \mathcal{R}_T(\mathbf{x}) + \mathcal{R}_S(\mathbf{x}), \quad (7)$$

where \mathbf{y} is the fMRI data, and the Frobenius norm is $\|\mathbf{x}\|_F^2 = \sum_{i=1}^N |x(i, t)|^2$. The optimal solution is a compromise between the data fitness and the regularization penalties.

Temporal regularization

The rationale of the temporal regularization term $\mathcal{R}_T(x)$ is to exploit the sparsity of the innovation signal that can be derived from

the recovered activity-related signal. We further build on previous work (Karahanoglu et al., 2011) where we introduced a generalization of “total variation” (TV). As a brief reminder, the TV-norm of a 1-D signal $f(t)$ is defined as the ℓ_1 -norm of its derivative: $\|\Delta_D \{f\}(t)\|_1$. Since minimizing ℓ_1 favors sparse solutions, TV regularization leads to signals whose derivatives are sparse, which are piecewise-constant (Rudin et al., 1992). The generalization in Karahanoglu et al. (2011) introduced an additional linear differential operator in ℓ_1 -norm that can compensate for the presence of a linear system. In particular, we use the differential operator $\Delta_L = \Delta_D \Delta_{L_h}$ within the ℓ_1 -norm:

$$\mathcal{R}_T(\mathbf{x}) = \sum_{i=1}^V \lambda_1(i) \Delta_L \{\mathbf{x}(i, \cdot)\}_1, \quad (8)$$

where ℓ_1 -norm is defined as

$$\Delta_L \{\mathbf{x}(i, \cdot)\}_1 = \sum_{t=1}^N |\Delta_L \{x(i, t)\}|, \quad (9)$$

and $\lambda_1(i)$ is the regularization parameter for voxel i .

Spatial regularization

Since fMRI data has a large amount of spatial correlation, the spatial regularization $\mathcal{R}_S(\mathbf{x})$ promotes coherent activity within the same region. For the sake of illustration, regions are defined in this work based on an anatomical atlas; i.e., we assume M different parcels where R_k , $k = 1, \dots, M$, are the sets of voxels for each region. We then use a mixed $\ell_{(2,1)}$ -norm to express spatially coherent (smooth) activity inside a region and possibly crisp changes in activity across regions (Baritau et al., 2011; Yuan and Lin, 2006):

$$\mathcal{R}_S(\mathbf{x}) = \sum_{t=1}^N \lambda_2(t) \|\Delta_{Lap} \{\mathbf{x}(\cdot, t)\}\|_{(2,1)}, \quad (10)$$

where the $\ell_{(2,1)}$ -norm is defined as

$$\|\Delta_{Lap} \{\mathbf{x}(\cdot, t)\}\|_{(2,1)} = \sum_{k=1}^M \sqrt{\sum_{i \in R_k} \Delta_{Lap} \{x(i, t)\}^2}, \quad (11)$$

and Δ_{Lap} is the 3-D second-order difference (Laplacian) operator and $\lambda_2(t)$ is the regularization parameter for each timepoint. Spatial regularization will favor smooth activity patterns inside regions, but not across regions.

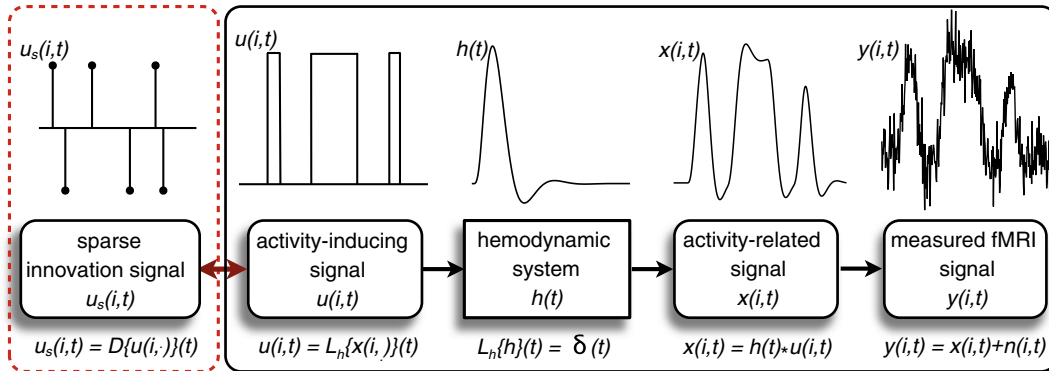


Fig. 1. fMRI signal model. Assuming that the activity-inducing signal is block-type, its derivative is the innovation signal, which will be sparse. The activity-related signal can be obtained by convolving the activity-inducing signal with the impulse response of the hemodynamic system. The activity-related signal is then further corrupted with noise and signal artifacts, and finally sampled at the fMRI temporal resolution (TR).

Optimization algorithm

We deploy the generalized forward–backward splitting algorithm (Raguet et al., 2012) to solve the optimization problem at hand. This iterative algorithm alternates between solving two cost functions:

$$\tilde{\mathbf{x}}_T = \arg \min_{\mathbf{x}} \frac{1}{2} \|\mathbf{y} - \mathbf{x}\|_2^2 + \mathcal{R}_T, \quad (12)$$

$$\tilde{\mathbf{x}}_S = \arg \min_{\mathbf{x}} \frac{1}{2} \|\mathbf{y} - \mathbf{x}\|_2^2 + \mathcal{R}_S. \quad (13)$$

In Fig. 2, we schematically outline the TA framework for fMRI data analysis. The extracted time courses are fed into this two step forward–backward splitting algorithm where a joint solution is achieved. Temporal regularization works with each voxel time course since the operator acts only in the temporal domain, whereas spatial regularization term works with each fMRI volume. We refer to Appendix A for further details about the algorithm.

Finally, we notice that the solution $\tilde{\mathbf{x}}$ will be the activity-related signal, however, we can apply the differential operator Δ_{L_h} to it to recover the activity-inducing signal as well.

Methods

Synthetic data

In order to validate the TA approach, we used a software phantom with $10 \times 10 \times 10$ voxels divided into four regions. The activity-inducing signal was fixed within a region, but different between regions. Two regions had spike-like activity-inducing signals: region 1

has spike trains with gradually increasing inter stimulus interval (ISI) from 1 to 12 s, region 2 has short events with duration uniformly distributed between [1,2]s. The other two regions have longer block-like activity (duration uniformly distributed between [1,...,15]s). The onset timings of the events have uniform distribution such that 12 and 6 events on average are generated in regions with spikes and blocks, respectively. A very short event is introduced into region 4 to test TA's robustness for short events in the middle of sustained events. All activity-inducing signals were sampled on a grid with temporal resolution (TR) of 1 s and had 200 timepoints. The activity-induced signals were then convolved with the HRF and corrupted with i.i.d. Gaussian noise such that the signal-to-noise ratio (SNR) was 1 dB. We define the SNR as the ratio of signal power to noise power in logarithmic scale as

$$\text{SNR} = 10 \log_{10} \left(\frac{\|\mathbf{x}\|^2}{\|\mathbf{y} - \mathbf{x}\|^2} \right).$$

Fig. 3 depicts the phantom and the associated time courses for each region.

In this example, temporal and spatial regularization operate in ideal settings. First, the temporal differential operator is perfectly matched with the HRF (see Supplementary for the parametrization of the operator Δ_{L_h}). Second, the spatial regularization uses the same regions as the generative model.

Experimental data

We further evaluated TA with experimental data acquired on 3 subjects during a sparse event-related paradigm where resting state periods

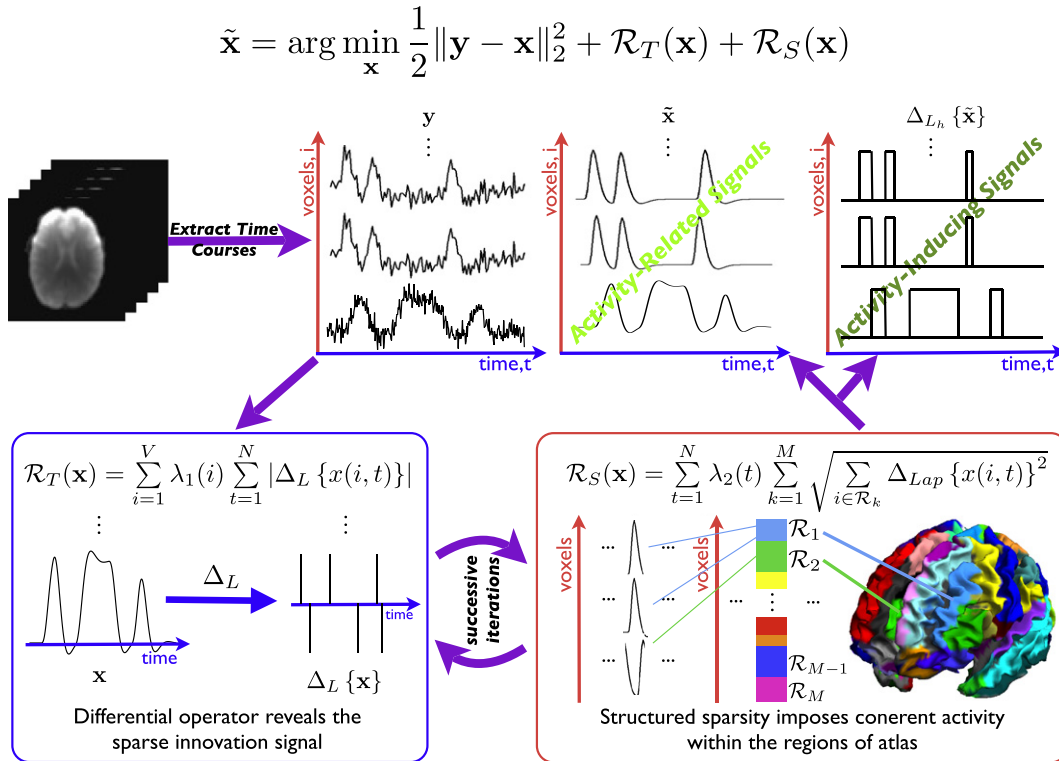


Fig. 2. Flowchart of TA. Successive regularization in temporal and spatial domains is applied to the noisy BOLD time courses. Temporal regularization (blue window) inverts the hemodynamic system by adopting a differential operator mimicking the inverse HRF and spatial regularization (red window) imposes smooth activations within the regions of an anatomical atlas. Finally, we obtain the activity-related signals and activity-inducing signals which reveal the neuronal-related activity.

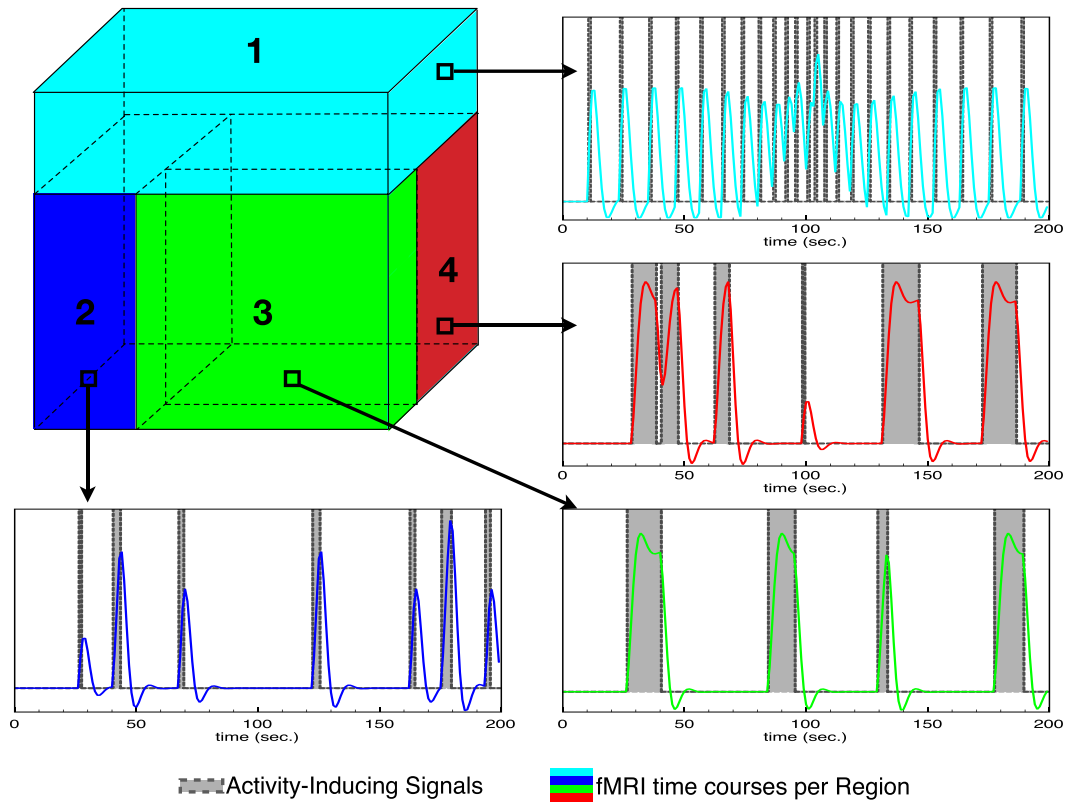


Fig. 3. The software phantom contains 4 regions in a cube that consists out of $10 \times 10 \times 10$ voxels. The first region (cyan, 300 voxels) was simulated as spike train with gradually increasing ISI from 1 s to 12 s and the second region (blue, 210 voxels) was simulated with random events with uniform duration in (Afshin-Pour et al., 2012; Aguirre et al., 1998) seconds. The third region (green, 245 voxels) and the fourth region (red, 245 voxels) were simulated with random events with uniform duration in (Afshin-Pour et al., 2012; Chang and Glover, 2010) seconds. A very short event is inserted into region 4 (around 100 s). The time resolution was chosen as $TR = 1$ s. The activity-inducing signals (in gray) were convolved with HRF to obtain the BOLD activity for each region. Each voxel time series was then corrupted with i.i.d. Gaussian noise such that voxel time series had SNR of 1 dB.

were disrupted by 10 visual stimuli of 8 Hz flickering checkerboard of duration 1 s with onsets randomly chosen following a uniform distribution during the duration of the run. When no visual stimuli were presented, subjects were instructed to maintain visual fixation on a cross in the center of the screen. The experiment was conducted in a Siemens TIM Trio 3 T MR scanner with a 32-channel head coil. The fMRI data comprised $N = 160$ (subjects 1 and 3) and $N = 190$ (subject 2) T_2 -weighted gradient echo-planar volumes ($TR/TE/FA$: 2 s/30 ms/85°, voxel size: $3.25 \times 3.25 \times 3.5$ mm³, matrix = 64×64). A T1-weighted MPRAGE anatomical image was also acquired during the MR session (192 slices, $TR/TE/FA$: 1.9 s/2.32 ms/9°, voxel size: $0.45 \times 0.45 \times 0.9$ mm³, matrix = 512×512).

The preprocessing steps included realignment of the datasets to the first scan of each subject and then spatial smoothing with a Gaussian smoother ($FWHM = 5$ mm). The spatial smoothing is not an essential step since TA also includes spatial regularization. However, the temporal regularization parameter is tuned for each voxel with respect to the (estimated) noise level. Therefore, spatial smoothing provides a spatially smoother estimate (less variations) of the noise level. Both steps were performed in the functional space of the subjects using SPM8 (FIL, UCL, UK). The anatomical AAL atlas (90 regions without the cerebellum) was mapped onto each subject's functional space using the IBASPM toolbox (Alemán-Gómez et al., 2006; Tzourio-Mazoyer et al., 2002). The voxels' time courses labeled within the atlas were detrended using a first-degree polynomial (i.e., linear trend) and slow oscillations (i.e., DCT basis function up to cut-off frequency of 1/250 Hz), and finally scaled to have unit variance. As before, the temporal differential operator was chosen from linear inverse of the Balloon model (Khalidov et al., 2011) (see supplementary

section). The regularization parameters need to provide a compromise between data fitness and regularization cost. We calibrated the temporal regularization parameter such that the residual noise level converged to the pre-estimated noise level of the data fit, where pre-estimated noise level is derived from the median absolute deviation of fine-scale wavelet coefficients (Daubechies, order 3). Then, for each iteration n , we update the temporal regularization parameter $\lambda_1(i)^n$ (Algorithm 2, step 10) similar to Chambolle (2004):

$$\lambda_1(i)^{n+1} = \frac{N\tilde{\sigma}(i)}{\frac{1}{2}\|\mathbf{y}(i, \cdot) - \mathbf{x}(i, \cdot)\|_2^2} \lambda_1(i)^n.$$

Spatial regularization parameter was empirically selected to be 5, which seemed to compensate well between temporal and spatial priors. The overall computation time for one dataset was around 5 h using a Linux cluster with Matlab (version 7.9).

After applying TA, we obtain three spatiotemporal datasets per subject: (1) the innovation signal $u_s(i, t)$; (2) the activity-inducing signal $u(i, t)$; (3) the activity-related signal $x(i, t)$. The innovation signal is the driver of the others, which can be derived through linear convolutions. To summarize the rich amount of information available in these datasets, we computed the average of the activity-inducing signals within each anatomical region, and then obtained the Spearman correlation matrix between the averaged time courses. Correlations were Fisher z-transformed and fed into a Ward's hierarchical clustering algorithm implemented in Matlab (Mathworks, Natick, MA, version (7.9) function linkage.m) to reveal the network structure contained in the activity-inducing signals. We selected two different levels to cut the dendrogram in order to show the evolution of clusters with

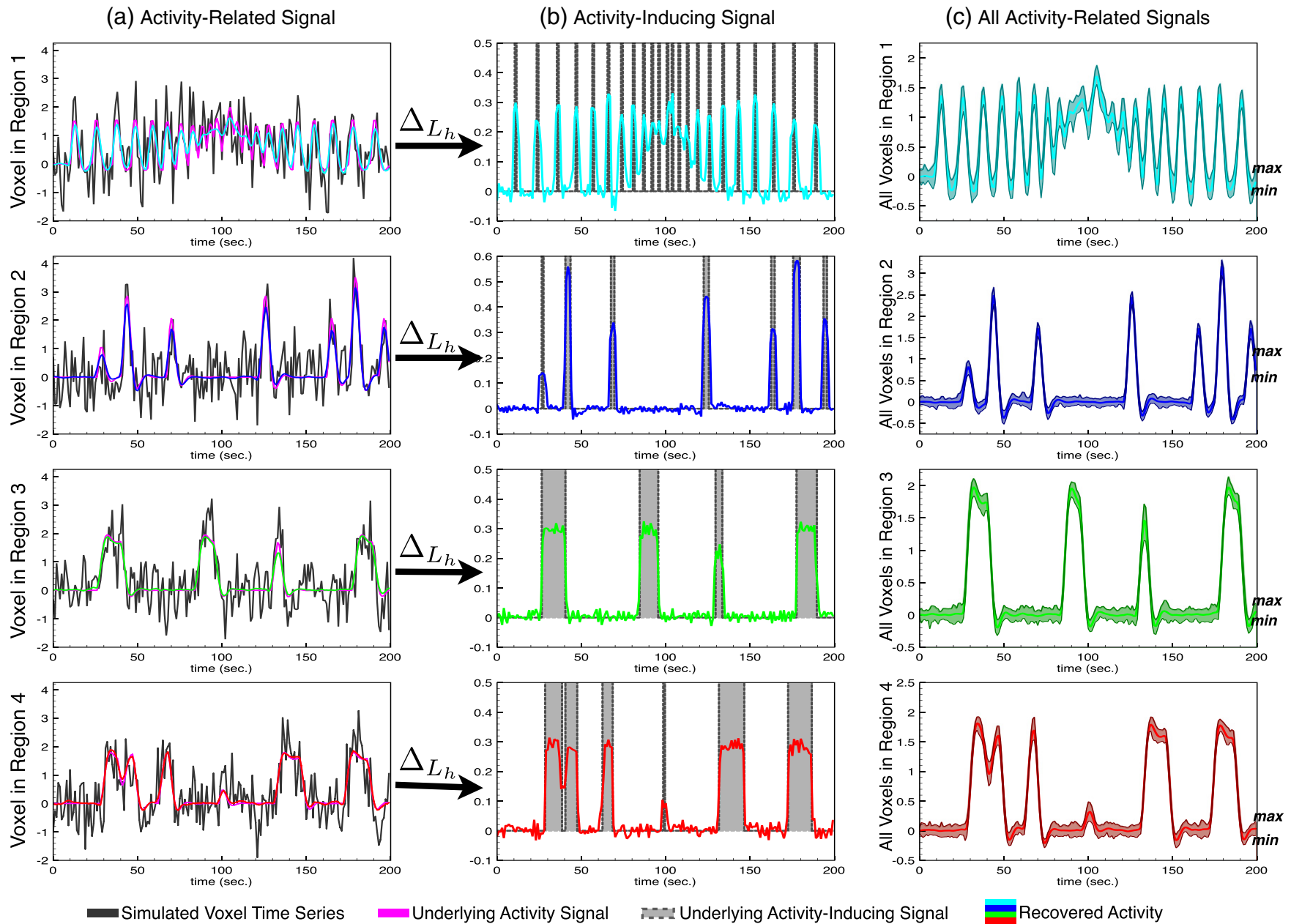


Fig. 4. Results for the software phantom. The left column (a) shows simulated noisy data (black), underlying BOLD signals (magenta), and recovered activity-related signals of a random voxel in each region (cyan, blue, red, green, respectively). The middle column (b) shows the underlying activity-inducing signal (gray) and the associated recovered activity signals. Finally, the right column (c) shows all recovered activity-related signals from the software phantom indicated by their mean, maximum, minimum per region. Small deviations within each region are observed.

respect to the inconsistency criterion that measures the deviation in each cluster.

Results

Synthetic data

In Figs. 4(a) and (b), we show, for randomly selected voxels in the four different regions, the activity-related and activity-induced signals, respectively. The recovered activity-inducing signals match very closely with the ground-truth underlying activity with no prior information on the timing or duration of the simulated events. In Fig. 4 (first row), we observe for region 1 that TA can resolve for events with ISI down to 2 TRs. The signal model is able to successfully recover different types of activity-inducing signals; i.e., short spike-like and long block-like stimuli, especially the short event in region 4 is well detected, but with lower amplitude. In Fig. 4(c), the variation of activity-related signals per region is summarized within the shaded area. Despite the relatively high level of noise in the simulated time courses, the recovered activity-inducing signals have small deviations across the voxels within each region. We refer the reader to Supplementary Figs. S.1 and S.2 to see the results without spatial regularization. Additional simulations were performed to a) simulate errors in the information provided by

the spatial template where TA demonstrated robustness against this model mismatch except that very brief activations may be undetected with large spatial discrepancies in Supplementary Fig. S.3, and b) vary the hemodynamic model used to generate the synthetic data (canonical HRF) with respect to the model used for the deconvolution (i.e. balloon model) which resulted in an expected time shift of the deconvolved signals due to the differences in the temporal characteristics of both models, but without altering a successful recovery of the underlying activations in Supplementary Fig. S.4.

Experimental data

After applying TA, we averaged the region-averaged correlation matrix of the three subjects and obtained the dendrogram in Fig. 5 as a result of hierarchical clustering. We extracted functionally distinct clusters at coarse (high) and detailed (low) levels. In other words, going down from the highest level in the dendrogram (whole brain) the consistency in the hierarchy is gradually incremented until a first group of clusters is defined (high-level), further increasing the consistency splits the clusters into subclusters which are meaningful segregations (low-level). At the high-level hierarchy, the brain is segregated into 9 global clusters (represented in different colors in the dendrogram); at the low-level hierarchy, 17 local networks (subclusters pinned from (1a) to (9) in

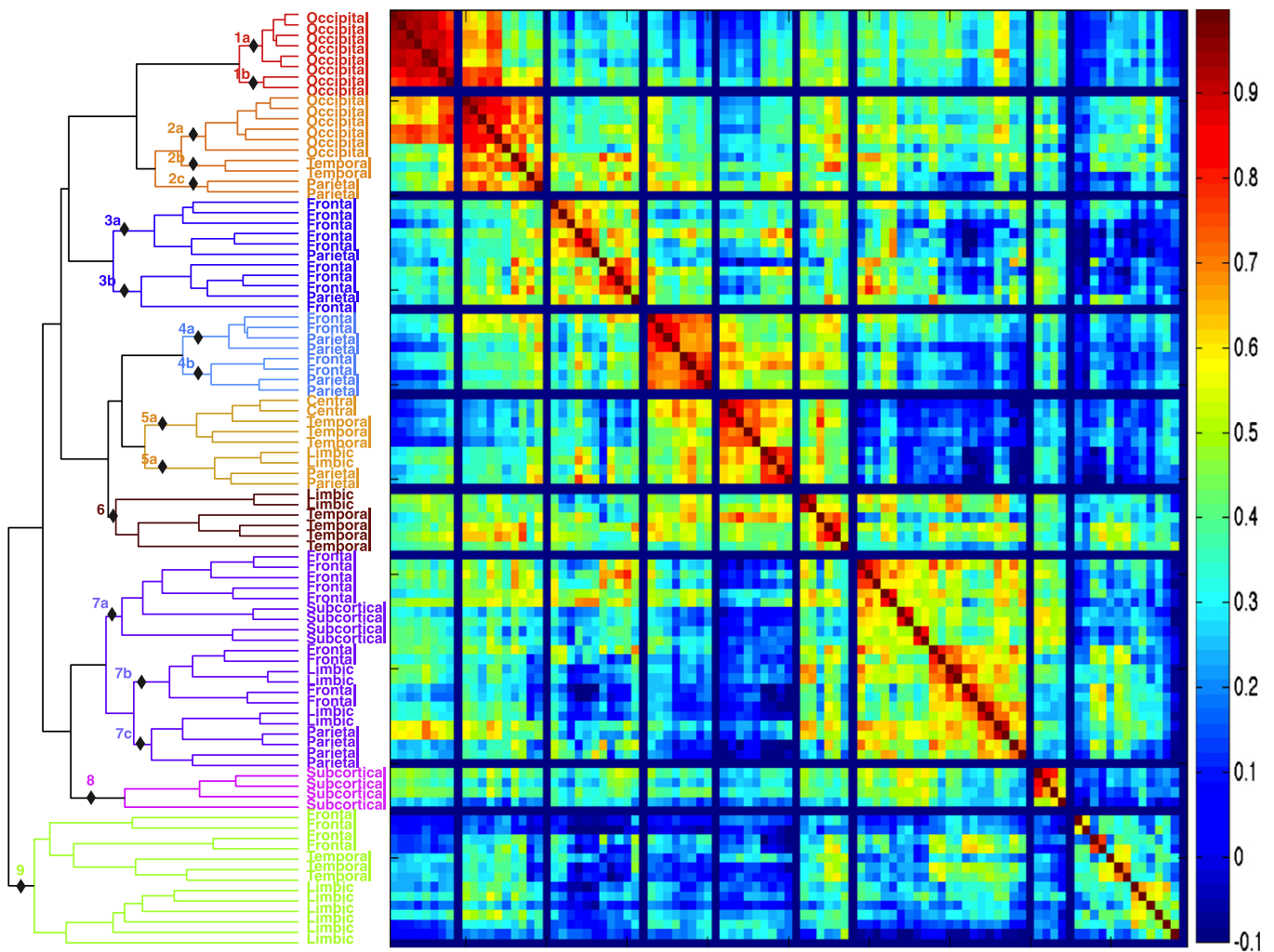


Fig. 5. Correlation matrix and corresponding clusters for TA activity-inducing signal. The dendrogram that reflects the hierarchical organization is shown on the left. Each color is described with the regions correspond to a different cluster in high level clusters (total 9 clusters) which is evaluated via inconsistency measure. Note that low level clusters (marked with black pins in the dendrogram from (1a) to (9) subdivide the clusters resulting 17 clusters. The anatomical descriptions in the clusters are detailed in Table 1.

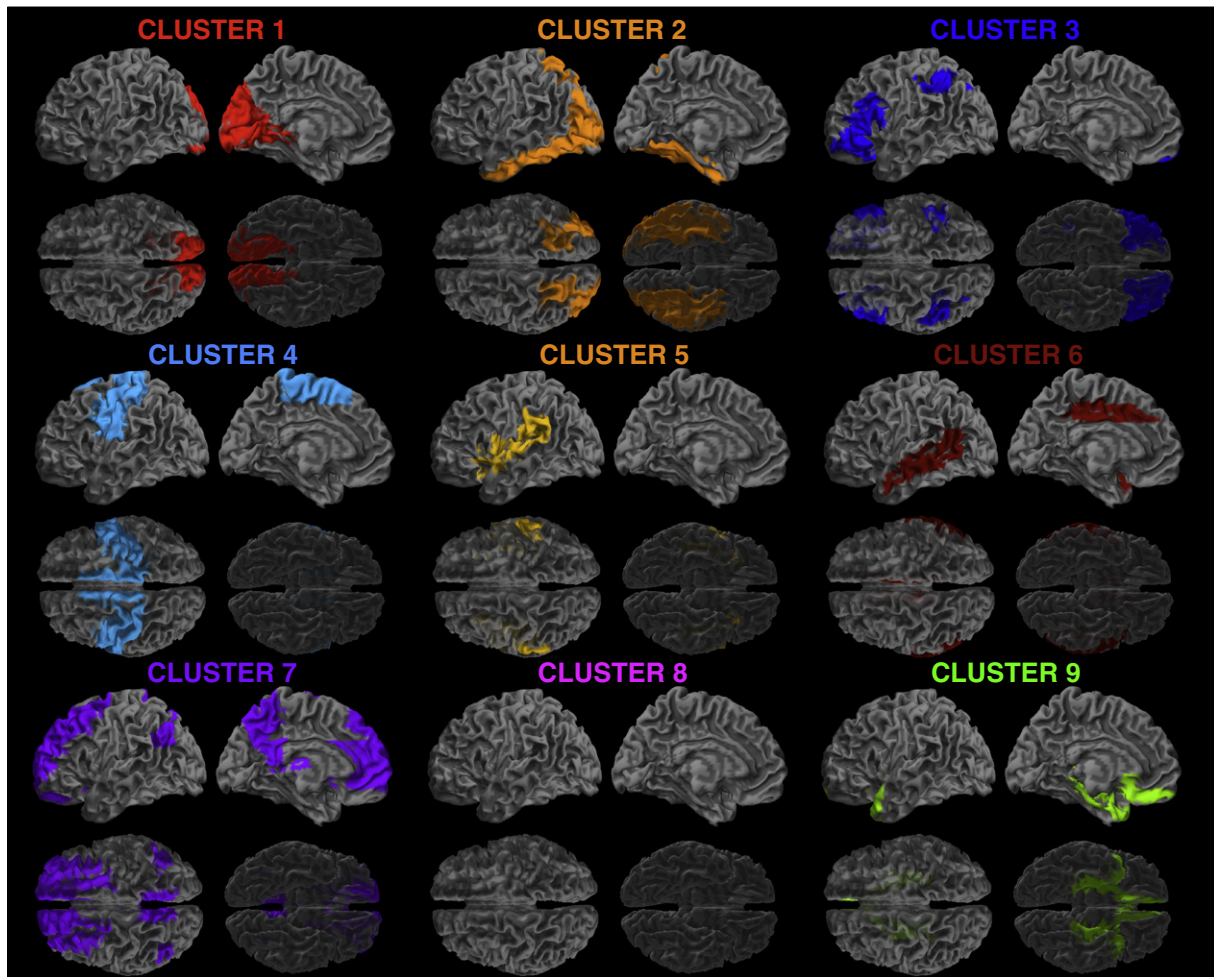


Fig. 6. Brain maps for the 9 high-level hierarchy clusters viewed from sagittal left (top left), sagittal cross-section in the middle (top right), top view (bottom left) and bottom view (bottom right). The regions are generated using anatomical atlas in MNI space corresponding to the anatomical descriptions in Table 1. We recover the activity-related networks; i.e., primary and late visual networks in clusters 1 and 2, respectively. Additionally, the fronto-parietal network (cluster 3), motor and somatosensory regions (cluster 4) and auditory network (cluster 5) as well as the default mode network (cluster 7), subcorticals (cluster 8) and limbic system (cluster 9) are observed. The clusters are nicely organized bilaterally.

the dendrogram) are revealed. Fig. 6 illustrates the high-level networks on the anatomical atlas. The extended anatomical descriptions in each (sub)cluster are listed in Table 1. We detail these clusters according to the order of the dendrogram.

The visual network makes up the first and second clusters, which is expected due to the stimulation and its strong coherence in resting state. Cluster 1 contains primary visual areas such as calcarine fissure, lingual gyrus and cuneus. Cluster 2 includes higher level visual areas extending towards ventral and dorsal visual pathways, inferior temporal gyrus and superior parietal lobule, which are subclusters 2b and 2c, respectively. In Fig. 7 (bottom right), the region-averaged activity-inducing signal in the visual network confirms that the timing of the visual stimuli (red bars) is well recovered without any prior knowledge. Cluster 3 reveals a fronto-parietal network extending bilateral middle frontal gyrus, inferior frontal gyrus and inferior parietal lobule, which mimics the dorsal attention network (Fox et al., 2006) and involves in attentional mechanisms, especially for “salient and unattended events” (Corbetta and Shulman, 2002). Subclusters 3a and 3b represent the right and left lateralized fronto-parietal regions similar to Beckmann et al. (2005) and Damoiseaux et al. (2006), respectively.

Cluster 4 reveals sensory-motor areas including primary motor cortex, primary somatosensory cortex as subcluster 4a, and supplementary

motor areas as subcluster 4b. Cluster 5 maps the auditory network where speech and language processing occur, including the Heschl gyrus, superior temporal gyrus (Wernicke's area) and inferior frontal gyrus. Cluster 6 involves bilateral midcingulate cortex, middle temporal gyrus as well as the right superior temporal gyrus. Cluster 7 consists of superior and middle frontal gyrus, anterior–posterior cingulate cortex (PCC) representing the default mode network (DMN) including thalamus (Luca et al., 2006). The hierarchical clustering suggests that cluster 7 is segregated into its anterior (7a, 7b) and posterior (7c) components, which are known to be part of saliency and executive control networks (Fox et al., 2006; Seeley et al., 2007), respectively. Similar subdivisions of the DMN have also been reported recently using real-time fMRI neurofeedback (Van De Ville et al., 2012). Subcortical regions, putamen and pallidum, are engaged in cluster 8 bilaterally. Cluster 9 involves bilateral limbic regions, parahippocampal gyrus, hippocampus and amygdala, as well as olfactory bulb, gyrus rectus and temporal poles. For the sake of comparison the results of the same correlation and hierarchical clustering analysis on the original detrended data, i.e. without processing with our TA method, are shown in Supplementary Figs. S.5, S.6 and Table S.1. We observe that visual, motor and auditory networks are also identified, however, they are given different preferences in dendrogram (auditory network is cluster 8 instead on 5, motor is 7 instead of 4). In addition, the most prominent right–left lateralized fronto-parietal

Table 1

The list of regions in the clusters. The clustering algorithm delineates 9 and 17 clusters in the high and low-level hierarchies (also presented in dendrogram in Fig. 5). The first two clusters correspond to the visual networks. Note that cluster 3 (fronto-parietal network) is subdivided into its right (3a) and left (3b) compartments in the higher hierarchy. Likewise, cluster 7 (default mode) is divided into its anterior (7a, 7b) and posterior (7c) components.

Cluster	Lobe	Anatomical description
1a	Occipital	Calcarine Fissure Left
	Occipital	Calcarine Fissure Right
	Occipital	Lingual Gyrus Left
	Occipital	Lingual Gyrus Right
	Occipital	Cuneus Left
1b	Occipital	Cuneus Right
	Occipital	Superior Occipital Gyrus Left
2a	Occipital	Superior Occipital Gyrus Right
	Occipital	Middle Occipital Gyrus Left
2b	Occipital	Inferior Occipital Gyrus Left
	Occipital	Middle Occipital Gyrus Right
	Occipital	Fusiform Gyrus Left
	Occipital	Fusiform Gyrus Right
	Occipital	Inferior Occipital Gyrus Right
2c	Temporal	Inferior Temporal Gyrus Left
	Temporal	Inferior Temporal Gyrus Right
3a	Parietal	Superior Parietal Gyrus Left
	Parietal	Superior Parietal Gyrus Right
3b	Frontal	Superior Frontal Gyrus (Orbital) Right
	Frontal	Inferior Frontal Gyrus (Orbital) Right
	Frontal	Middle Frontal Gyrus (Orbital) Right
	Frontal	Inferior Frontal Gyrus (Opercular) Right
	Frontal	Inferior Frontal Gyrus (Triangular) Right
	Parietal	Inferior Parietal Gyrus Right
	Parietal	Middle Frontal Gyrus (Orbital) Left
4a	Frontal	Inferior Frontal Gyrus (Opercular) Left
	Frontal	Inferior Frontal Gyrus (Triangular) Left
	Parietal	Inferior Parietal Gyrus Left
	Frontal	Inferior Frontal Gyrus (Orbital) Left
	Frontal	Precentral Gyrus Left
4b	Frontal	Precentral Gyrus Right
	Parietal	Postcentral Gyrus Left
	Parietal	Postcentral Gyrus Right
5a	Frontal	Supplementary Motor Area Left
	Frontal	Supplementary Motor Area Right
	Parietal	Paracentral Lobule Left
	Parietal	Paracentral Lobule Right
	Central	Rolandic Operculum Left
5b	Central	Rolandic Operculum Right
	Temporal	Superior Temporal Gyrus Left
	Temporal	Heschl Gyrus Right
6	Temporal	Heschl Gyrus Left
	Limbic	Insula Left
	Limbic	Insula Right
	Parietal	SupraMarginal Gyrus Left
	Parietal	SupraMarginal Gyrus Right
7a	Limbic	Medial Cingulate Cortex Left
	Limbic	Medial Cingulate Cortex Right
	Temporal	Superior Temporal Gyrus Right
	Temporal	Middle Temporal Gyrus Left
	Temporal	Middle Temporal Gyrus Right
	Temporal	Temporal Pole (Superior) Right
	Frontal	Superior Frontal Gyrus (Orbital) Left
7b	Frontal	Middle Frontal Gyrus Left
	Frontal	Superior Frontal Gyrus (Dorsolateral) Left
	Frontal	Superior Frontal Gyrus (Dorsolateral) Right
	Frontal	Middle Frontal Gyrus Right
	Subcortical	Caudate Nucleus Left
	Subcortical	Caudate Nucleus Right
	Subcortical	Thalamus Left
7c	Subcortical	Thalamus Right
	Frontal	Superior Frontal Gyrus (Medial) Left
	Frontal	Superior Frontal Gyrus (Medial) Right
	Limbic	Anterior Cingulate Cortex Left
	Limbic	Anterior Cingulate Cortex Right
7c	Frontal	Superior Frontal Gyrus (Medial–Orbital) Left
	Frontal	Superior Frontal Gyrus (Medial–Orbital) Right
	Limbic	Posterior Cingulate Cortex Left
Parietal	Posterior Cingulate Cortex Right	
Parietal	Precuneus Left	

Table 1 (continued)

Cluster	Lobe	Anatomical description
8	Parietal	Precuneus Right
	Parietal	Angular Gyrus Left
	Parietal	Angular Gyrus Right
	Subcortical	Putamen Left
	Subcortical	Pallidum Left
9	Subcortical	Putamen Right
	Subcortical	Pallidum Right
	Frontal	Olfactory Cortex Left
	Frontal	Olfactory Cortex Right
	Frontal	Gyrus Rectus Left
	Frontal	Gyrus Rectus Right
	Temporal	Temporal Pole (Superior) Left
9	Temporal	Temporal Pole (Middle) Left
	Temporal	Temporal Pole (Middle) Right
	Limbic	Hippocampus Left
	Limbic	ParaHippocampal Gyrus Left
	Limbic	Hippocampus Right
	Limbic	ParaHippocampal Gyrus Right
	Limbic	Amygdala Left
Limbic	Amygdala Right	

network and anterior–posterior segregation of the default mode network is lost in the hierarchy.

Having identified coherent networks through clustering of the activity-inducing signals, we can try to represent the dynamics for brain regions revealed by TA. Fig. 7 depicts the average activity-inducing signals rearranged according to the clusters. While the stimulus timings are well discovered mainly in the clusters corresponding to visual areas, we observe spontaneous activity in the visual network which does not correspond to visual stimuli (e.g., subject 2, cluster 1, around 300 s).

Fig. 8 shows the dynamic activity-inducing maps of subject 2. Two time courses are picked randomly from cuneus and PCC in order to track the temporal evolution of the task-related and spontaneous events. The positive and negative activations in PCC lagging the stimulus reflect the alternating structure of functional reorganization in the brain.

Finally, since the temporal prior of TA favors block-like activity-inducing signals, we evaluate the average block-length per region as the 4th quartile of the activity duration, see Fig. 9. It can be seen that regions in the visual clusters have relatively shorter average activity than other regions.

Discussion

Methodological implications

We have proposed TA to deconvolve fMRI data based on hemodynamic and anatomical properties of the brain. The objective is formulated as a minimization problem where the convex cost function contains sparsity-inducing regularization terms in both the temporal and spatial dimensions. The optimization is performed using a state-of-the-art generalized forward–backward scheme. The variational cost function is formulated assuming uncorrelated noise structure, nevertheless, an autoregressive noise model can be easily integrated into the framework. The colored noise should be whitened based on the estimated covariance of the residuals, which then leads to a weighted ℓ_2 -norm for the data-term of the cost function.

While TA does not require any timing information, it is not completely model-free neither—the three main underlying assumptions are: (1) the operator to invert the hemodynamic system; (2) the sparse innovation signal that leads to block-type activity; (3) the anatomical atlas for spatial regularization. In the following discussion, we

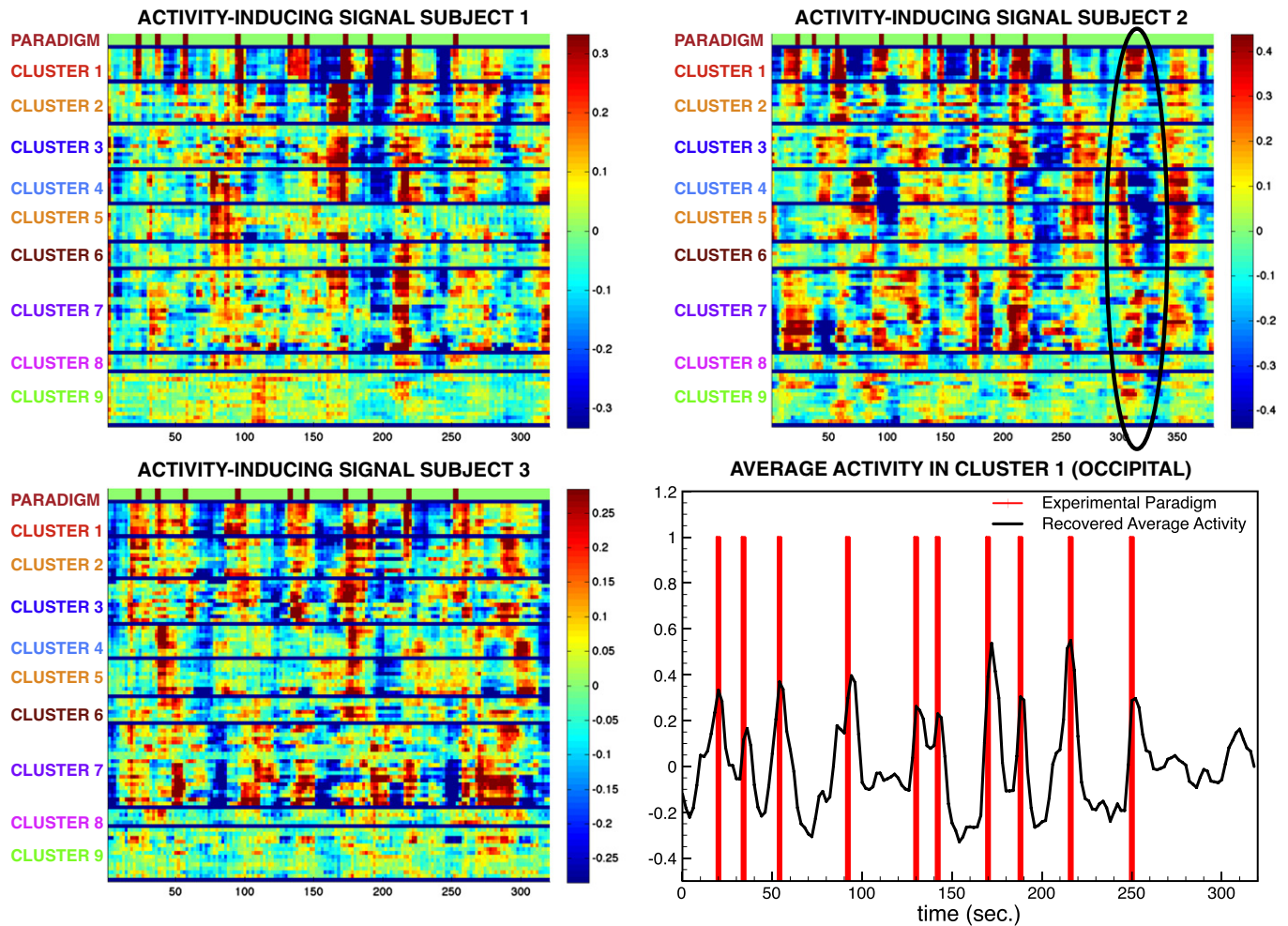


Fig. 7. Activity-inducing signals per region and subject. Clearly, the activity-inducing signal in the visual regions (clusters 1 and 2) follows the visual paradigm closely. Moreover, we observe the intrinsic brain activity, for example, a spontaneous event (in black contour) occurs around 300 s in subject 2 which is followed by negative activation in clusters 4–7 (posterior default mode network). The average activity in the occipital lobe (bottom right) matches with the visual stimulation.

first highlight each of these assumptions and then the related work in the field.

Hemodynamics

We use the first-order Volterra kernel of the non-linear balloon model to characterize the differential operator that relates activity-inducing signal with fMRI BOLD. It is well known that this model is limited since high variability of HRF within and across subjects has been observed (Handwerker et al., 2004). Additionally, the HRF might be different for task-related versus spontaneous activity. Nevertheless, the current approach seems to be a reasonable starting point. It can be replaced by other more sophisticated models, such as the one by Sotero and Trujillo-Barreto (2007), which accounts for the difference in excitatory and inhibitory neuronal activity. Another possible extension is to fine-tune the HRF operator (e.g., per region or even voxel) by alternating between the optimization of the activity-inducing signal and the HRF parameters (4 zeros and 1 pole) to obtain the region/voxel specific HRF similar to parcel based HRF estimation methods in Chaari et al. (2013) and Vincent et al. (2010).

Sparse innovation signal

The core principle of TA consists of defining an innovation signal that should be sparse if the experimentally-evoked or spontaneous brain activity has an underlying activity-inducing signal that is block-type. We have demonstrated that this is a strong prior, which allows us to recover paradigm-related activity without any timing information, yet also

without constraints on the block length of the activity-inducing signal. Besides, this signal model goes beyond the traditional Fourier viewpoint; i.e., for signals with the same regularization cost, the Fourier spectrum can vary from “simple” (e.g., periodic timing) to “complex” depending on the relative timing of the events.

Here we have imposed the ℓ_1 -norm for the temporal prior, which is the most common sparsity-pursuing norm. Although the Bayesian interpretation of the variational framework is beyond the scope of this paper, we would like to mention that stochastic processes generating sparse (innovation) signals are an active research in signal processing and applied statistics. In particular, the continuous-domain definition of this type of “noise” requires proper analytical tools (Unser et al., 2011). An intriguing research avenue for generative models of activity-inducing signals, is the choice of the stochastic process to generate the innovation signal.

Anatomical atlas

Many methods for fMRI data analysis rely upon Gaussian smoothing to improve the SNR. The spatial regularization of TA adds a structured-sparsity constraint by favoring smoothing within anatomical regions, but sparsity across. This is obtained by a mixed norm expressed on the Laplacian of the volume. For our settings of the spatial regularization parameter, we obtain more consistent activity-inducing signals within the regions, but at the same time the signals are not necessarily the same (e.g., see Fig. 8).

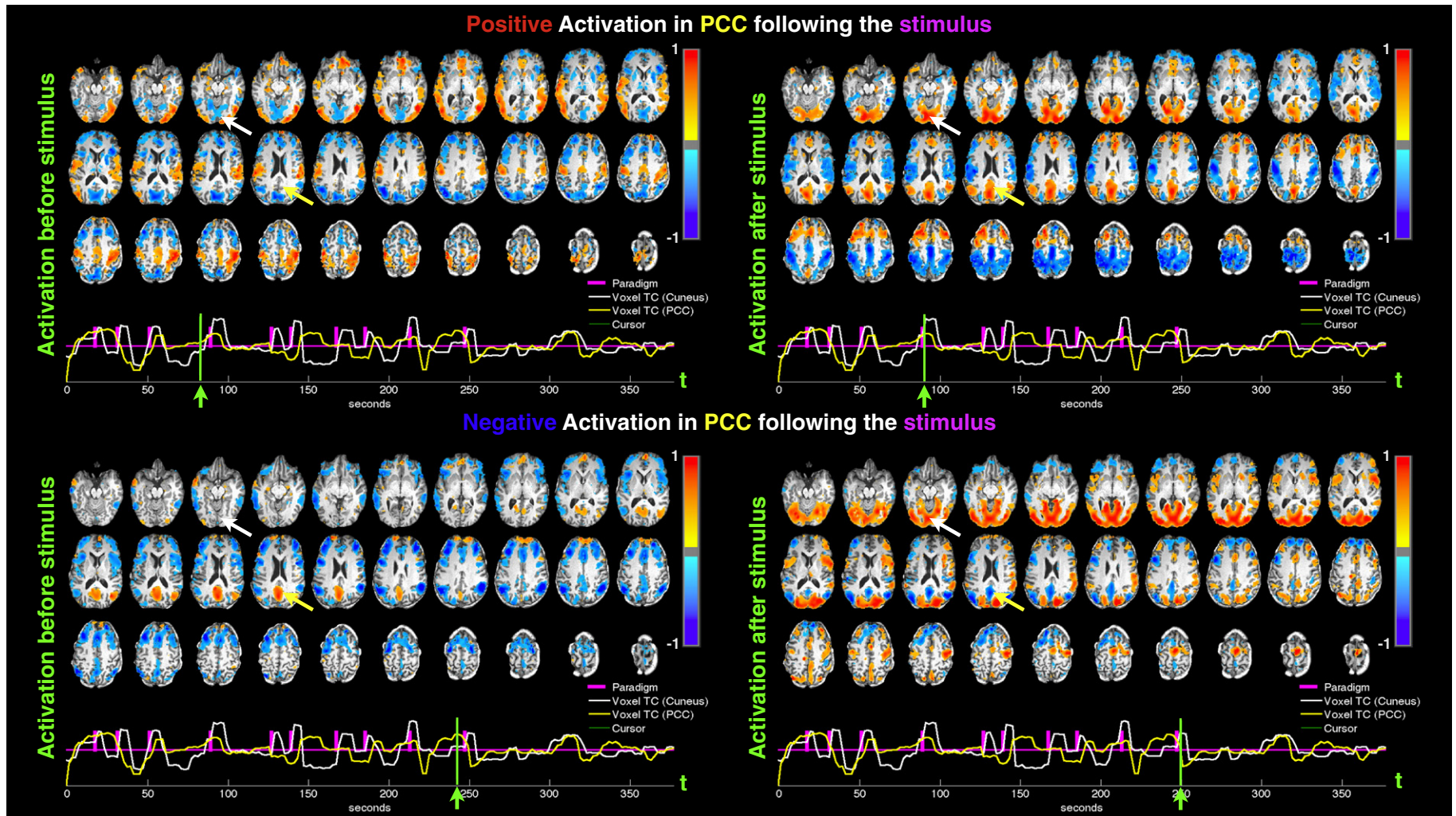


Fig. 8. Reconstructed dynamic activity-inducing maps for subject 2 (see supplementary for the movie). Time courses from cuneus and PCC are plotted in white and yellow, respectively. The stimulus time course is shown in magenta. In the first frame, the activity maps are illustrated for two instances (top row around 90 s and bottom row around 250 s). The left and right columns show the activation maps before and after the stimulus, respectively. PCC lags the stimuli with positive (top row) or negative response (bottom row).

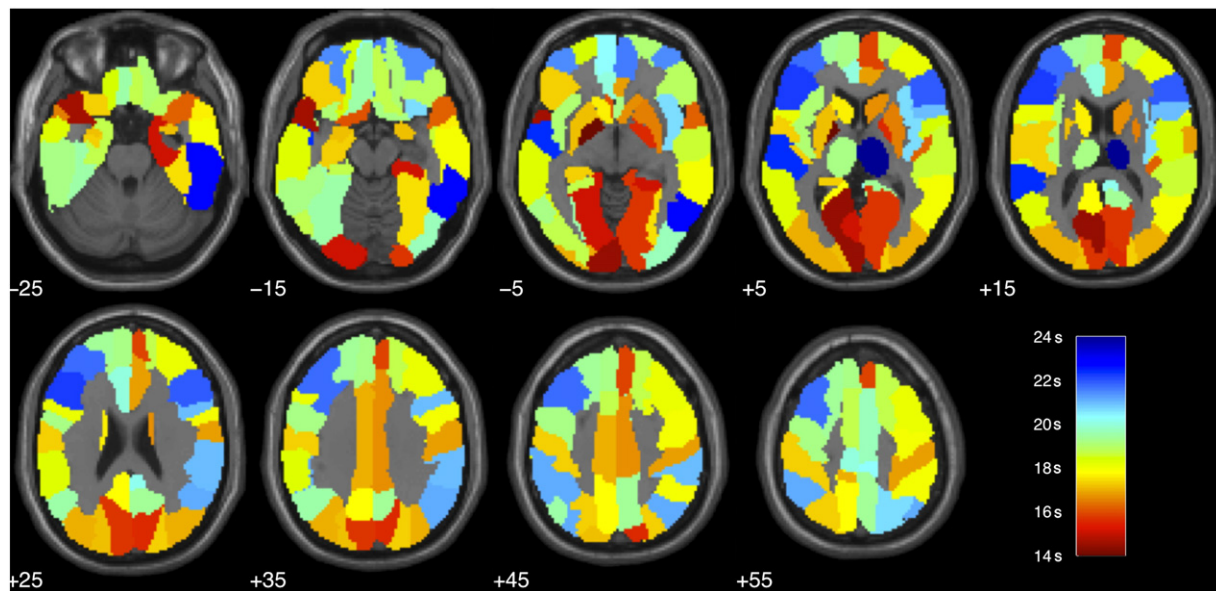


Fig. 9. Average block length for the region-averaged activity-inducing signals. Regions in the visual network have relatively shorter average activity than other regions.

The anatomical atlas is still a limitation of TA since it is not specific to the subject. An interesting option would be to make the atlas data-adaptive. For example, the regions could be re-estimated after a couple of iterations of the algorithm (Bellec et al., 2010) or joint estimation of HRF and HRF-defined region identification could be adapted into TA (Chaari et al., 2012). Another option would be to apply TA to the time courses recovered from a mode-decomposition (e.g., ICA).

Related methods

We mention in more detail three other deconvolution methods that have deployed similar ideas to recover activity-inducing signals (Gaudes et al., 2011, 2013; Hernandez-Garcia and Ulfarsson, 2011; Khalidov et al., 2011). All these methods operate solely in the temporal domain, and do not add spatial information into their framework. Also, the temporal prior of these methods is a “synthesis prior”, which means that an explicit dictionary of atoms is built. In Khalidov et al. (2011), wavelets (“activelets”) tailored to the hemodynamic system are designed and used to decompose BOLD fMRI signals that should ideally lead to sparse activelet coefficients if the activity is spike-like. In Gaudes et al. (2011) and Gaudes et al. (2013), a dictionary with all possible shifts of the canonical HRF is constructed to recover spike-like activity. In Hernandez-Garcia and Ulfarsson (2011), the dictionary combines spike-like and block-like atoms with two different regularization terms, which raises the issue how to choose the regularization parameters to adapt to the ratio of spikes and blocks for each voxel. Alternatively, the temporal regularization of TA uses an “analysis prior”, which is much more flexible (e.g., blocks of different lengths do not need to be specified explicitly, which would be prohibitive).

Dynamics of activity-inducing maps

When visualizing the activity-inducing signals obtained by TA on dynamic brain maps, we can easily recognize the presence and onsets of the visual stimuli. However, the data is much richer and many spontaneous events are captured as well; e.g., strong activity in the visual network of subject 2 during the final resting period (Fig. 7). Interestingly, activity-inducing signals reveal some non-stationary relationships between the different brain regions; e.g., as can be seen from Fig. 8, the signals from PCC and visual cortex are sometimes negatively correlated, and sometimes positively. This non-stationary behavior is also suggested in a time-frequency coherence analysis of fMRI (Chang and Glover, 2010).

Moreover, a recent work by Smith et al. (2012) exploits temporal and spatial ICA on high resolution data to reveal the temporally-independent and spatially overlapping activity maps called “temporal functional modes”. The authors show that different networks share common subcomponents of each other, that is, one brain region does not necessarily belong to a distinct functional network.

Hierarchical clustering

As a post-processing step, clustering the activity-inducing signals obtained by TA allows us to get a better understanding of the data. We obtain functionally plausible networks (many bilateral) reflecting both task-related and spontaneous activity. It is somewhat intriguing that activity of both task-related and resting-state networks is so well captured. While it is known that spontaneous activity continues during task and configurations similar to resting-state networks are formed (Fox and Raichle, 2007), it also means that our model for block-like activity-inducing signals is well suited for both types of activity. This raises the interesting hypothesis for future studies whether resting-state activity is rather block-like (with long durations on average) versus sinusoidal fluctuations (with low frequency).

The block model for the activity-inducing signal is also very flexible as it does not impose any duration. From the duration map, we clearly observe that the regions in the visual cortex have shorter duration whereas fronto-parietal regions have relatively the longest duration.

The quality of the clustering (i.e., high correlation within and low correlation between clusters) is higher for the results obtained on TA-regularized activity-inducing signals; see Figs. 5 and 6 and compare with Supplementary Figs. S.5 and S.6. For non-TA processed signals, we obtain higher variation in the large clusters (especially clusters 3 and 4). Visual, motor and auditory networks match most with the results of TA, however, their rank in dendrogram is different (auditory network is cluster 8 instead on 5, motor is 7 instead of 4). Moreover, the most prominent right–left lateralized fronto-parietal network and anterior–posterior segregation of the default mode networks is lost in the hierarchy.

Conclusion

We have introduced TA, a new analysis tool that essentially aims at revealing the activity-inducing driver of BOLD fMRI. Using synthetic and experimental data, we evaluated TA’s ability to recover the underlying activity without timing information of cognitive

tasks or stimuli, as well as the intrinsic brain activity. The method is formulated within the variational framework and solved using a convex optimization scheme where temporal and spatial priors are handled iteratively. This study allows for exploring brain dynamics while retaining both hemodynamics and spatial coherence in the brain and hence constitute a good candidate to reveal the both task-related and spontaneous activity. The application of TA to uncover dynamic activity-inducing maps could also be a powerful tool prior to more advanced clustering techniques.

Acknowledgments

This work was supported in part by the Swiss National Science Foundation under grant PP00P2-123438 and in part by the Center for Biomedical Imaging (CIBM) of the Geneva-Lausanne Universities & Hospitals and the EPFL.

Appendix A. Algorithm

The generalized forward–backward splitting (Raguet et al., 2012) aims to minimize a convex cost function that consists of a quadratic data fitting term and multiple regularization terms. The solution is obtained by incorporating the solutions of each convex regularization problem separately which are defined as

$$\tilde{\mathbf{x}}_T = \arg \min_{\mathbf{x}} \frac{1}{2} \|\mathbf{y} - \mathbf{x}\|_F^2 + \sum_{i=1}^V \lambda_1(i) \|\Delta_L \{\mathbf{x}(i, \cdot)\}\|_1 = \text{prox}_{\mathcal{R}_T}(\mathbf{y}), \quad (\text{A.1})$$

$$\tilde{\mathbf{x}}_S = \arg \min_{\mathbf{x}} \frac{1}{2} \|\mathbf{y} - \mathbf{x}\|_F^2 + \sum_{t=1}^N \lambda_2(t) \|\Delta_S \{\mathbf{x}(\cdot, t)\}\|_{(2,1)} = \text{prox}_{\mathcal{R}_S}(\mathbf{y}). \quad (\text{A.2})$$

The solutions of the above problems correspond to proximal maps to the temporal and spatial regularization terms in Eqs. (8) and (10) (Combettes and Wajs, 2005). These variational formulations are nonquadratic; i.e., direct solutions do not exist, and hence iterative algorithms should be employed. Forward–backward splitting method (Combettes and Wajs, 2005) applies on functionals when the regularization term is a non-smooth convex function and the data-term is a smooth differentiable function with a finite Lipschitz constant. In the forward step, a gradient descent is applied for data-term and in the backward step, proximal map is computed for the regularization function. We exploit the dual definitions of ℓ_1 -norm and $\ell_{2,1}$ -norm (Baritau et al., 2011; Chambolle, 2004) to find the solution of each proximal map, and employ “fast gradient projection” (FGP) (Beck and Teboulle, 2009) to achieve faster convergence. Here, we formulate the generalized algorithm (detailed in Raguet et al., 2012).

Algorithm 1. Spatio-temporal regularization $\tilde{\mathbf{x}} = \arg \min_{\mathbf{x}} \frac{1}{2} \|\mathbf{y} - \mathbf{x}\|_F^2 + \mathcal{R}_T(\mathbf{x}) + \mathcal{R}_S(\mathbf{x})$

INPUTS: Noisy fmri data \mathbf{y} ,
 1: Initialize: $n \leftarrow 1$, $\hat{\mathbf{x}}_T^0 = 0$, $\hat{\mathbf{x}}_S^0 = 0$, $\tilde{\mathbf{x}}^0 = 0$,
 2: **Repeat**
 3: Solve for temporal prior: $\hat{\mathbf{x}}_T^{n+1} = \hat{\mathbf{x}}_T^n + \text{prox}_{\mathcal{R}_T}(\hat{\mathbf{x}}_T^n - \hat{\mathbf{x}}_S^n + \mathbf{y}) - \tilde{\mathbf{x}}^n$
 4: Solve for spatial prior: $\hat{\mathbf{x}}_S^{n+1} = \hat{\mathbf{x}}_S^n + \text{prox}_{\mathcal{R}_S}(\hat{\mathbf{x}}_T^{n+1} - \hat{\mathbf{x}}_S^n + \mathbf{y}) - \tilde{\mathbf{x}}^n$
 5: Update $\tilde{\mathbf{x}}^{n+1} = \hat{\mathbf{x}}_T^{n+1}/2 + \hat{\mathbf{x}}_S^{n+1}/2$
 6: $n \leftarrow n + 1$
 7: **until** convergence or number of maximum iterations are reached.

Let's consider the temporal term where the cost function is

$$C_T(\mathbf{x}(i, \cdot)) = \mathcal{F}(\mathbf{x}) + \sum_{i=1}^V \lambda_1(i) \|\Delta_L \{\mathbf{x}(i, \cdot)\}\|_1, \quad (\text{A.3})$$

where the smooth data-term is $\mathcal{F}(\mathbf{x}) = \frac{1}{2} \|\mathbf{y} - \mathbf{x}\|_F^2$. The regularization term is ℓ_1 -norm; i.e., a non-smooth function. Let's call that the minimizer of C_T is \mathbf{x}^* , then

$$\partial C_T(\mathbf{x}^*) \in \nabla \mathcal{F}(\mathbf{x}^*) + \partial \mathcal{R}_T(\mathbf{x}^*) = 0, \quad \mathbf{x}^* - \mu \nabla \mathcal{F}(\mathbf{x}^*) \in (\mathbf{I} + \mu \partial \mathcal{R}_T)(\mathbf{x}^*), \\ \mathbf{x}^* = (\mathbf{I} + \mu \partial \mathcal{R}_T)^{-1}(\mathbf{x}^* - \mu \nabla \mathcal{F}(\mathbf{x}^*)) = \text{prox}_{\mu \mathcal{R}_T}(\mathbf{x}^* - \mu \nabla \mathcal{F}(\mathbf{x}^*)).$$

The iterative solution consists of solving for forward term, i.e., $\mathbf{x}_{\text{forward}} = \mathbf{x}^n - \mu \nabla \mathcal{F}(\mathbf{x}^*)$ and backward term, i.e., $\mathbf{x}^{n+1} = \text{prox}_{\mu \mathcal{R}_T}(\mathbf{x}_{\text{forward}})$. Specifically, forward solution in Eq. (A.3) is straightforward and minimizer depends only on the proximal map with $\mathbf{x}_{\text{forward}} = \mathbf{y}$ and $\mathbf{x}^* = \text{prox}_{\mathcal{R}_T}(\mathbf{y})$.

There is no direct analytical solution for minimizing the primal problem in Eq. (A.3). Instead the dual formulation provides an easier interpretation. The dual-norm is defined as

$$\|\Delta_L \{\mathbf{x}(i, \cdot)\}\|_1 = \max_{\|\mathbf{p}\|_1 < 1} \langle \Delta_L \{\mathbf{x}(i, \cdot)\}, \mathbf{p}(i, \cdot) \rangle.$$

Then,

$$\min_{\mathbf{x}} C_T = \min_{\mathbf{x}} \max_{\|\mathbf{p}\|_1 \leq 2} \frac{1}{2} \|\mathbf{y} - \mathbf{x}\|_F^2 + \sum_{i=1}^V \lambda_1(i) \langle \mathbf{x}(i, \cdot), \Delta_L^T \{\mathbf{p}(i, \cdot)\} \rangle,$$

since function is convex in \mathbf{x} concave in \mathbf{p} we interchange min–max and at the saddle point the primal and dual formulations lead to the same solution

$$= \max_{\|\mathbf{p}\|_1 \leq 1} \min_{\mathbf{x}} \frac{1}{2} \|\mathbf{y} - \mathbf{x}\|_F^2 + \sum_{i=1}^V \lambda_1(i) \langle \mathbf{x}(i, \cdot), \Delta_L^T \{\mathbf{p}(i, \cdot)\} \rangle,$$

where Δ_L^T corresponds to the adjoint of Δ_L , i.e., $\Delta_L^T[t] = \Delta_L[-t]$.

This algorithm has a fast convergence due to steps 7 and 8 (Beck and Teboulle, 2009; Combettes and Wajs, 2005; Raguet et al., 2012) and a robust update of the regularization parameter λ (Chambolle, 2004). The algorithm for spatial regularization can be achieved similarly for each volume t exploiting the dual-norm of $\ell_{(2,1)}$ -norm into the formulation instead of ℓ_1 -norm (Baritau et al., 2011). Discrete implementation of operator Δ_L can be efficiently done by constructing stable (causal and anti-causal) filters. We refer to Karahanoglu et al. (2011) for implementation details and filter coefficients.

Algorithm 2. Temporal regularization $\text{prox}_{\mathcal{R}_T}(\mathbf{y}) = \arg \min_{\mathbf{x}} \frac{1}{2} \|\mathbf{y} - \mathbf{x}\|_F^2 + \sum_{i=1}^V \lambda_1(i) \|\Delta_L \{\mathbf{x}\}\|_1$

INPUTS: Noisy fmri data \mathbf{y} , Differential operator Δ_L , Noise estimate $\tilde{\sigma}$ that is, median absolute deviation of the second order wavelet coefficients of \mathbf{y} and spectral norm $c > \sup_{\omega} |\hat{\Delta}_L(e^{i\omega})|^2 = \sup_{\omega} \prod_{m=1}^4 \frac{1 - e^{i\omega} e^{-j\omega}}{1 - e^{i\omega} e^{-j\omega}} |2$
 1: **for** each voxel $i = 1$ to V **do**
 2: $n \leftarrow 1$
 3: Initialize: $k^1 = 1$, $\mathbf{p}^0 = 0, \mathbf{x}(i, \cdot)^1 = 0, \mathbf{v}^1 = 0$ and $\lambda_1(i)^1 = \tilde{\sigma}(i)$
 4: **repeat**
 5: Update $\mathbf{p}^n = \Delta_L \{\mathbf{y}(i, \cdot)\} / (\lambda_1(i)^n c) + (I - \Delta_L^T/c) \{\mathbf{v}^n\}$
 6: Update $\mathbf{p}^n = \text{clip}(\mathbf{p}^n)$ where $\text{clip}(\cdot)$ denotes the element-wise clipping function,
 $\text{clip}(\cdot) = \text{sign}(\cdot) \min(|\cdot|, 1)$,
 7: Update $k^{n+1} = \frac{1 + \sqrt{1 + 4(k^n)^2}}{2}$
 8: Update $\mathbf{v}^{n+1} = \mathbf{p}^n + \frac{k^n - 1}{k^{n+1}} (\mathbf{p}^n - \mathbf{p}^{n-1})$
 9: Set $\mathbf{x}(i, \cdot)^n = \mathbf{y}(i, \cdot) - \lambda_1(i)^n \Delta_L^T \{\mathbf{p}^n\}$,
 10: Update $\lambda_1(i)^{n+1} = \frac{N \tilde{\sigma}(i)}{\|\mathbf{y}(i, \cdot) - \mathbf{x}(i, \cdot)\|_2} \lambda_1(i)^n$,
 11: $n \leftarrow n + 1$
 12: **until** convergence or number of maximum iterations are reached.
 13: **end for**

Appendix B. Supplementary data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.neuroimage.2013.01.067>.

References

- Afshin-Pour, B., Hossein-Zadeh, G.A., Strother, S.C., Soltanian-Zadeh, H., 2012. Enhancing reproducibility of fMRI statistical maps using generalized canonical correlation analysis in NPAIRS framework. *Neuroimage* 60, 1970–1981.
- Aguirre, G., Zarahn, E., D'Esposito, M., 1998. The variability of human, BOLD hemodynamic responses. *Neuroimage* 8, 360–369.
- Alemán-Gómez, Y., Melie-García, L., Valdés-Hernandez, P., 2006. IBASPM: toolbox for automatic parcellation of brain structures. 12th Annual Meeting of the Organization for Human Brain Mapping, 27.
- Baritoux, J.C., Hassler, K., Bucher, M., Sanyal, S., Unser, M., 2011. Sparsity-driven reconstruction for FDOT with anatomical priors. *IEEE Trans. Med. Imaging* 30, 1143–1153.
- Baumgartner, R., Ryner, L., Richter, W., Summers, R., Jarmasz, M., Somorjai, R., 2000. Comparison of two exploratory data analysis methods for fMRI: fuzzy clustering vs. principal component analysis. *Magn. Reson. Imaging* 18, 89–94.
- Beck, A., Tebouille, M., 2009. Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems. *IEEE Trans. Signal Process.* 18, 2419–2434.
- Beckmann, C.F., Smith, S.M., 2004. Probabilistic independent component analysis for fMRI. *IEEE Trans. Med. Imaging* 23, 137–152.
- Beckmann, C.F., DeLuca, M., Devlin, J.T., Smith, S.M., 2005. Investigations into resting-state connectivity using independent component analysis. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 360, 1001–1013.
- Bellec, P., Rosa-Neto, P., Lyttelton, O.C., Benali, H., Evans, A.C., 2010. Multi-level bootstrap analysis of stable clusters in resting-state fMRI. *Neuroimage* 51, 1126–1139.
- Biswal, B., Yetkin, F.Z., Haughton, V., Hyde, J., 1995. Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magn. Reson. Med.* 34, 537–541.
- Calhoun, V., Adali, T., 2006. Unmixing fMRI with independent component analysis. *IEEE Eng. Med. Biol. Mag.* 25, 79–90.
- Chaari, L., Forbes, F., Vincent, T., Ciuciu, P., 2012. Adaptive hemodynamic-informed parcellation of fMRI data in a variational joint detection estimation framework. 15th Proc. MICCAI, LNCS. Springer Verlag, pp. 180–188.
- Chaari, L., Vincent, T., Forbes, F., Dojat, M., Ciuciu, P., 2013. Fast joint detection-estimation of evoked brain activity in event-related fMRI using a variational approach. *IEEE Trans. Med. Imaging*. <http://dx.doi.org/10.1109/TMI.2012.2225636>.
- Chambolle, A., 2004. An algorithm for total variation minimization and applications. *J. Math. Imaging Vision* 20, 89–97.
- Chang, C., Glover, G.H., 2010. Time-frequency dynamics of resting-state brain connectivity measured with fMRI. *Neuroimage* 50, 81–98.
- Cole, D.M., Smith, S.M., Beckmann, C.F., 2010. Advances and pitfalls in the analysis and interpretation of resting-state fMRI data. *Front. Syst. Neurosci.* 4.
- Combettes, P., Wajs, V., 2005. Signal recovery by proximal forward-backward splitting. *Multiscale Model. Simul.* 4, 1168–1200.
- Corbetta, M., Shulman, G.L., 2002. Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* 3, 201–215.
- Damoiseaux, J.S., Rombouts, S.A.R.B., Barkhof, F., Scheltens, P., Stam, C.J., Smith, S.M., Beckmann, C.F., 2006. Consistent resting-state networks across healthy subjects. *Proc. Natl. Acad. Sci.* 103, 13848–13853.
- Evans, J., Todd, R., Taylor, M., Strother, S., 2010. Group specific optimisation of fMRI processing steps for child and adult data. *Neuroimage* 50, 479–490.
- Flandin, G., Penny, W.D., 2007. Bayesian fMRI data analysis with sparse spatial basis function priors. *Neuroimage* 34, 1108–1125.
- Fox, M.D., Raichle, M.E., 2007. Spontaneous fluctuations in brain activity observed with fMRI. *Nat. Rev. Neurosci.* 8, 700–711.
- Fox, M.D., Corbetta, M., Snyder, A.Z., Vincent, J.L., Raichle, M.E., 2006. Spontaneous neuronal activity distinguishes human dorsal and ventral attention systems. *Proc. Natl. Acad. Sci.* 103, 10046–10051.
- Friston, K.J., Fletcher, P., Josephs, O., Holmes, A., Rugg, M.D., Turner, R., 1998. Event-related fMRI: characterizing differential responses. *Neuroimage* 7, 30–40.
- Friston, K.J., Mechelli, A., Turner, R., Price, C.J., 2000. Nonlinear responses in fMRI: the balloon model, Volterra kernels, and other hemodynamics. *Neuroimage* 12, 466–477.
- Friston, K., Trujillo-Barreto, N., Daunizeau, J., 2008. DEM: a variational treatment of dynamic systems. *Neuroimage* 41, 849–885.
- Friston, K.J., Stephan, K., Daunizeau, J., 2010. Generalised filtering. *Math. Probl. Eng.* 2010. <http://dx.doi.org/10.1155/2010/621670>.
- Gaodes, C.C., Petridou, N., Dryden, I.L., Bai, L., Francis, S.T., Gowland, P.A., 2011. Detection and characterization of single-trial fMRI BOLD responses: paradigm free mapping. *Hum. Brain Mapp.* 32, 1400–1418.
- Gaodes, C.C., Petridou, N., Francis, S.T., Dryden, I.L., Gowland, P.A., 2013. Paradigm free mapping with sparse regression automatically detects single-trial fMRI BOLD responses. *Hum. Brain Mapp.* 39, 501–518.
- Gitelman, D., Penny, W., Ashburner, J., Friston, K., 2003. Modeling regional and psychophysiological interactions in fMRI: the importance of hemodynamic deconvolution. *Neuroimage* 19, 200–207.
- Glover, G.H., 1999. Deconvolution of impulse response in event-related BOLD fMRI. *Neuroimage* 9, 416–429.
- Gusnard, D.A., Raichle, M.E., 2001. Searching for a baseline: functional imaging and the resting human brain. *Nat. Rev. Neurosci.* 2, 685–694.
- Handwerker, D.A., Ollinger, J.M., D'Esposito, M., 2004. Variation of BOLD hemodynamic responses across subjects and brain regions and their effects on statistical analyses. *Neuroimage* 21, 1639–1651.
- Harrison, L., Penny, W., Flandin, G., Ruff, C., Weiskopf, N., Friston, K., 2008. Graph-partitioned spatial priors for functional magnetic resonance images. *Neuroimage* 43, 694–707.
- Havlicek, M., Friston, K.J., Jan, J., Brazdil, M., Calhoun, V.D., 2011. Dynamic modeling of neuronal responses in fMRI using cubature Kalman filtering. *Neuroimage* 56, 2109–2128.
- Hernandez-García, L., Ulfarsson, M.O., 2011. Neuronal event detection in fMRI time series using iterative deconvolution techniques. *Magn. Reson. Imaging* 29, 353–364.
- Karahanoğlu, F.I., Bayram, I., Van De Ville, D., 2011. A signal processing approach to generalized 1-D total variation. *IEEE Trans. Signal Process.* 59, 5265–5274.
- Khalidov, I., Van De Ville, D., Fadili, J., Unser, M., 2007. Activelets and sparsity: a new way to detect brain activation from fMRI data. Proceedings of the SPIE Conference on Mathematical Imaging: Wavelet XII, San Diego CA, USA, pp. 1–8.
- Khalidov, I., Fadili, J., Lazeyras, F., Van De Ville, D., Unser, M., 2011. Activelets: Wavelets for sparse representation of hemodynamic responses. *Signal Process.* 91, 2810–2821.
- Kirsch, A., 1996. An introduction to the mathematical theory of inverse problems, 2nd edition. *Applied Mathematical Sciences*, vol. 120. Springer.
- Liu, Y., Gao, J.H., Liu, H.L., Fox, P.T., 2000. The temporal response of the brain after eating revealed by fMRI. *Nature* 405, 1058–1062.
- Lopes, R., Lina, J., Fahoum, F., Gotman, J., 2012. Detection of epileptic activity in fmri without recording the EEG. *Neuroimage* 60, 1867–1879.
- Luca, M.D., Beckmann, C., Stefano, N.D., Matthews, P., Smith, S., 2006. fMRI resting state networks define distinct modes of long-distance interactions in the human brain. *Neuroimage* 29, 1359–1367.
- Lund, T.E., Madsen, K.H., Sidaros, K., Luo, W.L., Nichols, T.E., 2006. Non-white noise in fMRI: does modelling have an impact? *Neuroimage* 29, 54–66.
- Makni, S., Idier, J., Vincent, T., Thirion, B., Dehaene-Lambertz, G., Ciuciu, P., 2008. A fully Bayesian approach to the parcel-based detection-estimation of brain activity in fMRI. *Neuroimage* 41, 941–969.
- Morgan, V.L., Li, Y., Abou-Khalil, B., Gore, J.C., 2008. Development of 2dTCA for the detection of irregular, transient BOLD activity. *Hum. Brain Mapp.* 29, 57–69.
- Petridou, N., Gaodes, C.C., Dryden, I.L., Francis, S.T., Gowland, P., 2012. Periods of rest in fMRI contain individual spontaneous events which are related to slowly fluctuating spontaneous activity. *Hum. Brain Mapp.* <http://dx.doi.org/10.1002/hbm.21513>.
- Raguet, H., Fadili, J., Peyr , G., 2012. Generalized forward-backward splitting. <http://arxiv.org/abs/1108.4404v3>.
- Raichle, M.E., 2006. The brain's dark energy. *Science* 314, 1249–1250.
- Riera, J.J., Watanabe, J., Kazuki, I., Naoki, M., Aubert, E., Ozaki, T., Kawashima, R., 2004. A state-space model of the hemodynamic approach: nonlinear filtering of BOLD signals. *Neuroimage* 21, 547–567.
- Rudin, L.I., Osher, S., Fatemi, E., 1992. Nonlinear total variation based noise removal algorithms. Proceedings of the Eleventh Annual International Conference of the Center for Nonlinear Studies on Experimental mathematics: Computational Issues in Nonlinear Science. Elsevier North-Holland, Inc., pp. 259–268.
- Seeley, W.W., Menon, V., Schatzberg, A.F., Keller, J., Glover, G.H., Kenna, H., Reiss, A.L., Greicius, M.D., 2007. Dissociable intrinsic connectivity networks for salience processing and executive control. *J. Neurosci.* 27, 2349–2356.
- Smith, M., Fahrmeir, L., 2007. Spatial Bayesian variable selection with application to functional magnetic resonance imaging. *J. Am. Stat. Assoc.* 102, 417–431.
- Smith, S.M., Miller, K.L., Moeller, S., Xu, J., Auerbach, E.J., Woolrich, M.W., Beckmann, C.F., Jenkinson, M., Andersson, J., Glasser, M.F., Essen, D.C.V., Feinberg, D.A., Yacoub, E.S., Ugurbil, K., 2012. Temporally-independent functional modes of spontaneous brain activity. *Proc. Natl. Acad. Sci.* 8, 3131–3136.
- Sotero, R.C., Trujillo-Barreto, N.J., 2007. Modelling the role of excitatory and inhibitory neuronal activity in the generation of the BOLD signal. *Neuroimage* 35, 149–165.
- Thirion, B., Flandin, G., Pinel, P., Roche, A., Ciuciu, P., Poline, J.B., 2006. Dealing with the shortcomings of spatial normalization: multi-subject parcellation of fMRI datasets. *Hum. Brain Mapp.* 27, 678–693.
- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Lhote, M., 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15, 273–289.
- Unser, M., Tafti, P.D., Sun, Q., 2011. A unified formulation of Gaussian vs. sparse stochastic processes part I: continuous-domain theory. <http://arxiv.org/abs/1108.6150v1>.
- Van De Ville, D., Jhooti, P., Haas, T., Kopel, R., Lovblad, K.O., Scheffler, K., Haller, S., 2012. Recovery and spatio-temporal segregation of default mode subnetworks. *Neuroimage* 63, 1175–1181.
- Vincent, T., Rissler, L., Ciuciu, P., 2010. Spatially adaptive mixture modeling for analysis of fMRI time series. *IEEE Trans. Med. Imaging* 29, 1059–1074.
- Yuan, M., Lin, Y., 2006. Model selection and estimation in regression with grouped variables. *J. R. Stat. Soc.* 68, 49–67.
- Zibulevsky, M., Elad, M., 2010. L1–L2 optimization in signal and image processing. *IEEE Signal Proc. Mag.* 27, 76–88.