

Subject to group comparison of fMRI data using group guided Dual Independent Component Analysis

Lorenz Berger

Abstract

Studies of functional MRI (fMRI) data increasingly concern estimation at the group level, especially differences in functional connectivity patterns between individual subjects and subject groups. Dual regression and Independent Component Analysis (ICA) based techniques have been used extensively in order to identify such patterns of functional connectivity for subject specific fMRI data. However the results can often be misinterpreted due to artifacts such as head motion which can contaminate the subject maps and potentially cause misdiagnosis. We present a two stage ICA approach implemented using Variational Bayes to solve this contamination problem and estimate subject specific maps more accurately. By working in a Bayesian framework we are able to use group map information to first compute individual subject time-courses, learn additional artifact time-courses, and then use these time-courses to estimate the final subject specific spatial maps. Comparison experiments against dual regression, the current state of the art algorithm demonstrate that the ‘Dual-ICA’ approach is able to obtain subject specific independent components that have less spatial contamination from artifacts.

Keywords: Functional Magnetic Resonance Imaging, Independent Component Analysis, Variational Bayes, Dual Regression

1. Introduction

1.1. Functional MRI

Functional magnetic resonance imaging is the application of MRI techniques to mapping brain activity, by imaging blood flow changes that are indirect, delayed measures of neurological activity. The subject is imaged using rapid pulse sequences that acquire an entire slice at once and reacquire the same slice repeatedly with a repeat time of around 3 seconds [1]. Thus a typical 10-minute scan will have around 200 time points for each of the roughly 1 million voxels in the 3D image. This fMRI data can then be analysed to find patterns of brain activity. These patterns are often referred to as spatial maps and have recently become of great interest to the neuroscientific community for probing neural mechanism and investigating diseases.

1.2. ICA for fMRI data

Independent Component Analysis (ICA) aims to extract features and structure from data that is assumed to be a linear combination of underlying independent components. ICA can be used to decompose the observed data, a two-dimensional (voxels $\times$ time) data matrix into a set of time-courses and associated spatial maps, which jointly describe the temporal and spatial characteristics of underlying hidden signals (components/maps). The main assumptions of ICA is that the underlying components are spatially independent, add linearly and are non-Gaussian. These assumptions have been shown to be correct for spatial maps produced by the activity-dependent sources of blood flow in the brain [2].

The ICA model presented in this work is a modified version of the Bayesian Linked-ICA model described in [3]. In [3] Linked-ICA is applied to multi-modal fMRI data for fusing different modality types together to automatically find patterns of related changes across multiple modalities. For our model we simply set the number of modalities to one and use the existing Bayesian framework to adjust the priors in such a way that we are able to incorporate group map information into our ICA estimation of subject specific maps. This Bayesian ICA model differs from standard methods like FastICA [4] in that it incorporates dimensionality reduction into the ICA method itself by the use of automatic relevance determination (ARD) priors on the components [5]. The model works on the full-dimensionality of the data and has an additive noise model, it also models an explicitly parametrized non-Gaussian source

model (in this case a Gaussian mixture model) instead of maximizing negentropy (as used in FastICA). Having an ARD in the model is useful, since we do not have to specify the number of components and can rely on the model to eliminate weak components. However this feature is also the main drawback of plain Bayesian ICA when applied to fMRI data, because brain maps that happened to have a weaker signal can get automatically removed. This is why we use group map information within the Dual-ICA procedure to ensure that all the maps are found.

1.3. Group ICA

Group ICA obtains subject specific maps by using a known group maps to guide the ICA. These known group maps are given as part of the input to the algorithm and can be found using various established methods. For example in [6] group maps from the fMRI data of multiple subjects are found by using a modified FastICA procedure to find the independent components (spatial maps). These maps are then converted to Z-score maps which depend on the amount of variability explained by the ICA decomposition, these are then thresholded to produced the final subject spatial maps. The final group maps can then be generated by averaging over several subjects. Once a set of group maps is available, subject specific maps can be computed using dual regression [7]. Differences between the subject and a group can then be used to help diagnose patients. For example, there has been a study by [8] which compares subjects associated with Alzheimer’s and other age related cognitive impairment with healthy subjects that are carriers of the APOE-e4 allele, a risk factor for developing Alzheimer’s later in life. It is found that Alzheimer patients show an increased amount of activation in the hippocampus which is associated with memory processing and lies within the ‘default mode network’ (Figure 1).

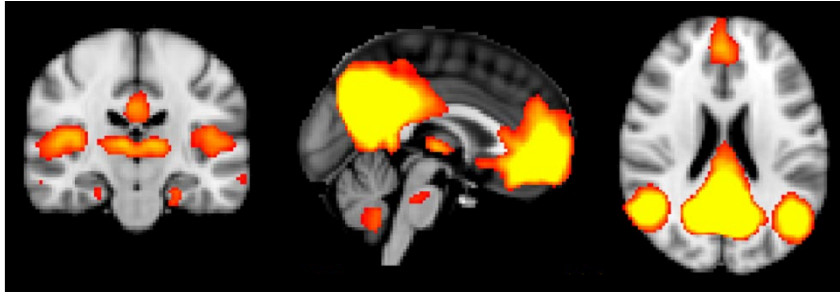


Figure 1: Shows the saggital, coronal and axial view of a spatial map representing the ‘default mode network’. This map is estimated from a group of 36 subjects using probabilistic ICA [6]. The images are taken from [8].

The problem with dual regression is that it does not always produce accurate subject specific maps, especially when artifacts and true maps are spatially or temporally correlated with each other. This correlation between artifacts and true maps causes the estimated maps to be a mixture of subject maps and artifacts, making it difficult for neuroscientists to identify the true spatial map. Throughout this work we will refer to this unwanted effect as contamination. By working within a Bayesian framework we are able to specify the priors in such a way that it allows us to incorporate the group map information to guide the subject specific ICA to find independent spatial maps as well as artifact maps, thus resulting in overall better map estimates that contain less contamination.

2. Bayesian Dual-ICA Model

The observed variables $\mathbf{Y}$ are the subject’s fMRI data which are made up of spatial maps $\mathbf{X}$ (voxels $\times$ components), time-courses $\mathbf{H}$ (components $\times$ time) and additive noise $\mathbf{E}$ such that

$$\mathbf{Y} = \mathbf{XH} + \mathbf{E}, \quad (1)$$

where $\mathbf{E}$ models Gaussian noise, with precision (inverse variance) λ :

$$\mathbf{E} \sim N(0, 1/\lambda), \quad (2)$$

$$\lambda \sim \text{Gamma}(\text{rms}(\mathbf{Y})^2 c_0, c_0). \quad (3)$$

Here rms is the root mean squared used to measure the magnitude of variability, for the constant c_0 a value of 10^{-3} has proven to be robust. The number of actual components (maps) is unknown but will be determined by initially assuming many components and then using an ARD prior [9] to eliminate components that are too weak and not significant. We therefore introduce the matrix $\mathbf{W}$ which can be interpreted as a weighting matrix for each component. The model now reads

$$\mathbf{Y} = \mathbf{X}\mathbf{W}\mathbf{H} + \mathbf{E}, \quad (4)$$

with $\mathbf{W}$ having the prior

$$\mathbf{W} \sim N(0, 1/\omega). \quad (5)$$

As one of the precisions ω for one of the components tends to ∞ it will effectively eliminate this component from that time point by forcing the corresponding row of $\mathbf{W}$ to be zero with very high precision.

2.1. Independent Spatial Maps

The driving force behind an ICA decomposition is that the data is derived from a number of statistically independent spatial sources; these are the spatial maps X_l for $l = 1 \dots L$. In order to find the non-Gaussianity of the individual sources we explicitly fit a non-Gaussian distribution to each source by assuming a particular functional form. This is the approach taken here, using an M-component Gaussian mixture model. The mixture model prior on the spatial maps can be expressed as

$$\begin{aligned} P(\mathbf{X}_{n,l} | \boldsymbol{\mu}, \boldsymbol{\beta}, \boldsymbol{\pi}) &= \sum_{m=1}^M P(q_n = m | \boldsymbol{\pi}_l) P(\mathbf{X}_{n,l} | m, \boldsymbol{\mu}_{l,m}, 1/\boldsymbol{\beta}_{l,m}) \\ &= \sum_{m=1}^M \pi_{l,m} N(\mathbf{X}_{n,l} | \boldsymbol{\mu}_{l,m}, 1/\boldsymbol{\beta}_{l,m}), \end{aligned} \quad (6)$$

where $\mathbf{X}_{n,l}$ is the n th voxel in the l th ICA component (spatial map), $\mu_{l,m}$, $\beta_{l,m}$ and $\pi_{l,m}$ are the means, precisions and component proportions respectively. Here $q_n = m$ is a hidden mixture membership variable that indicates which mixture component $\mathbf{X}_{n,l}$ was drawn from. For simplicity and practicality, we chose $M = 3$ as the number of mixture components which seems to work well. The priors for π come from a Dirichlet distribution

$$\boldsymbol{\pi} \sim \text{Dir}(1). \quad (7)$$

The priors for μ and β depend on which stage of the algorithm we are currently in, this will be discussed in section 3.

2.2. Variational Bayesian Inference

In Bayesian inference we aim to infer the posterior distributions for our parameters. Variational optimization can be applied to solve this inference problem, for a more complete explanation than the one given here we refer the reader to [10] and [11]. Let $\mathbf{Z}$ denote the set of all latent variables (maps and time-courses) and parameters, and let $\mathbf{X}$ denote the set of all observed variables (the fMRI data). Evaluating the full posterior distribution $P(\mathbf{Z} | \mathbf{X})$ is intractable so instead we choose a suitable distribution $Q(\mathbf{Z})$ which provides an approximation to the true posterior distribution. We can decompose the log marginal distribution as the lower bound $\mathcal{L}(Q)$ and the KL divergence $\text{KL}(Q || P)$ as

$$\ln P(\mathbf{X}) = \mathcal{L}(Q) + \text{KL}(Q || P), \quad (8)$$

where we have defined

$$\mathcal{L}(Q) = \int Q(\mathbf{Z}) \ln \left\{ \frac{P(\mathbf{Z}, \mathbf{X})}{Q(\mathbf{Z})} \right\} d\mathbf{Z}, \quad (9)$$

$$\text{KL}(Q || P) = - \int Q(\mathbf{Z}) \ln \left\{ \frac{P(\mathbf{Z} | \mathbf{X})}{Q(\mathbf{Z})} \right\} d\mathbf{Z}. \quad (10)$$

Our aim is to maximize the lower bound $\mathcal{L}(Q)$, which is equivalent to minimizing the KL divergence. For the choice of the variational Q -distribution we seek a choice that is simple enough to make our

computations tractable but at the same time gives enough flexibility to make the bound tight. Here we take the most common approach and use the mean field approximation, and the posterior distribution is factorized as

$$Q(\mathbf{Z}) = \prod_i Q(\mathbf{Z}_i) \quad (11)$$

$$= Q(\mathbf{H})Q(\beta)Q(\mu)Q(\pi)Q(\mathbf{W})Q(\omega)Q(\lambda) \prod_{l=1}^L Q(\mathbf{X}_l, \mathbf{q}_l). \quad (12)$$

This explicitly factorizes the spatial sources $\mathbf{X}$ over its separate components. This approximation is a reasonable one because the component sources are assumed to be statistically independent of one another. It can now be shown that the optimum form for each component of the posterior distribution is given by

$$Q(\mathbf{Z}_i) = \frac{\exp(\mathbb{E}_{k \neq i}[\ln P(\mathbf{X}, \mathbf{Z})])}{\int \exp(\mathbb{E}_{k \neq i}[\ln P(\mathbf{X}, \mathbf{Z})]) d\mathbf{Z}_j} \quad (13)$$

All the calculations for the expectations which are needed to obtain the parameter updates may be derived analytically and can be found in the appendix of [3]. The initialisation to this iterative procedure is specific to the stage of the algorithm, this is explained in the appendix.

3. Dual-ICA algorithm

Now that we have defined our model and shown how to solve it we can apply it to the previously described problem of finding subject specific maps $\mathbf{X}$ and time-courses $\mathbf{H}$ as well as any additional subject artifacts maps and time-courses $\hat{\mathbf{X}}$ and $\hat{\mathbf{H}}$ respectively.

3.1. Dual-ICA algorithm overview

First we ‘fix’ the known group maps as our subject maps by adjusting the priors in our model and then learn the corresponding time-courses. In addition to this we also use the ICA to learn extra artifact maps and time-courses. Then as a post-processing step to the first stage we inspect the time-courses found by the ICA and eliminate any duplicates. These duplicates correspond to differences between the subject and group maps, removing one of the duplicates will prevent the ICA splitting the subject map into two maps during the next stage. In the second and final stage we ‘fix’ all (both subject and artifact) time-courses previously found and set uninformative priors on the maps, allowing us to freely estimate the desired subject maps and artifact maps from the subject’s fMRI data.

All the coefficients used throughout this work have been chosen in line with [3] or were found experimentally. Because these hyperparameters are mostly extremely small or very large values, small changes in their value seem to have no significant consequence on the final result. However as part of future work the significance of these coefficients should be investigated to both show they are robust and to optimize the algorithm and improve results even further.

3.2. Stage 1, finding the time-courses

In this method we assume that we know the subject maps contain most of the group maps but do also have some map differences. In stage 1, we fix the subject maps to be the group maps and try to learn the corresponding time-courses, this is very similar to performing straight forward regression where the group maps are used as regressors. In addition to this we also learn extra maps and time-courses for artifacts, and subject maps that show differences between the subject and group maps. To implement this in our model we fix the prior of the spatial maps to be the known group maps

$$\mathbf{X}_f \sim N(\mathbf{G}, 10^{-6}), \quad (14)$$

and put an uninformative prior on the corresponding time-courses so they are ‘loose’ and are able to be estimated freely

$$\mathbf{H}_l \sim N(0, 10^6). \quad (15)$$

We also aim to find additional artifacts and map differences by adding extra spatial maps with uninformative priors

$$\dot{\mathbf{X}}_l \sim \sum_{m=1}^M \pi_m N(\boldsymbol{\mu}_m, 1/\boldsymbol{\beta}_m), \quad (16)$$

$$\boldsymbol{\mu} \sim N(0, 10^6), \quad (17)$$

$$\boldsymbol{\beta} \sim \text{Gamma}(10^3, 10^{-6}), \quad (18)$$

as well as uninformative priors for the corresponding time-courses

$$\dot{\mathbf{H}}_l \sim N(0, 1). \quad (19)$$

The priors on the ARD are

$$\boldsymbol{\omega}_f \sim \text{Gamma}(10^{-12}, 10^6), \quad (20)$$

$$\dot{\boldsymbol{\omega}}_l \sim \text{Gamma}(10^{12}, 10^{-2}). \quad (21)$$

We can also write this in matrix form as

$$\mathbf{Y} = [\mathbf{X}_f \quad \dot{\mathbf{X}}_l] \begin{bmatrix} \mathbf{W}_f & 0 \\ 0 & \dot{\mathbf{W}}_l \end{bmatrix} \begin{bmatrix} \mathbf{H}_l \\ \dot{\mathbf{H}}_l \end{bmatrix} + \mathbf{E}_l, \quad (22)$$

where the ICA has to estimate the whole right-hand side. All matrices with subscript f are denoted as fixed because they are unable to change from their initialisation, the other matrices with subscript l are loose and are able to be freely estimated.

3.3. Stage 1.1, getting rid of duplicate time-courses

From stage 1 we have learned subject time-courses $\mathbf{H}$ and additional time-courses $\dot{\mathbf{H}}$. As a post-processing step we inspect the additional time-courses $\dot{\mathbf{H}}$ and aim to identify time-courses which correspond to a spatial map difference between a subject and group map. These time-courses will be very similar to the time-courses learnt from previously fixed group maps that are different to the corresponding subject maps. We propose a simple threshold based method on the correlation between time-courses $\mathbf{H}$ and $\dot{\mathbf{H}}$

$$\text{corr}(\mathbf{H}_l, \dot{\mathbf{H}}_j) > \alpha \implies \text{matching time course found}, \quad (23)$$

where $l = [1 \dots L]$ are the different subject time-courses and $j = [1 \dots A]$ denote the additional time-courses, α is a threshold to decide whether two time-courses are similar enough to conclude that they correspond to the same spatial map, a value of $\alpha = 0.9$ seems to work well. Any time-courses in $\dot{\mathbf{H}}$ that have been found to have matching time-courses in $\mathbf{H}$ get removed from $\dot{\mathbf{H}}$. This is done to prevent the ICA from splitting the the one true map into two separate maps during stage 2.

3.4. Stage 2, finding the spatial maps

During stage 2 we fix all the previously found time-courses and then learn the corresponding spatial maps, this time these are the maps that best fit the subject's data. In matrix form we have

$$\mathbf{Y} = [\mathbf{X}_l \quad \dot{\mathbf{X}}_l] \begin{bmatrix} \mathbf{W}_f & 0 \\ 0 & \dot{\mathbf{W}}_f \end{bmatrix} \begin{bmatrix} \mathbf{H}_f \\ \dot{\mathbf{H}}_f \end{bmatrix} + \mathbf{E}_l. \quad (24)$$

Since we want to fix the time-courses we fix their priors as

$$\mathbf{H}_f \sim N(\mathbf{H}_{stage1}, 10^{-12}), \quad (25)$$

$$\dot{\mathbf{H}}_f \sim N(\dot{\mathbf{H}}_{stage1}, 10^{-12}). \quad (26)$$

The priors for the spatial maps are unconstrained (loose) so they can be freely estimated, with $\mathbf{X}_l$ having the same prior as $\dot{\mathbf{X}}_l$ in (16). The prior on $\dot{\boldsymbol{\omega}}_f$ is the same as $\boldsymbol{\omega}_f$ given in (20).

4. Results

4.1. Simulated fMRI

The simulated fMRI data that is used for our experiments was generated by mixing a set of spatial maps with distinctive time-courses using software downloaded from http://mlsp.umbc.edu/simulated_fmri_data.html. For this experiment we chose four distinct group maps that can be seen in Figure 2 and created an artificial subject containing four actual spatial maps, where two maps (maps 2 and 4) are the same as maps from the group maps and two are different (maps 1 and 3). The subject also contains an extra five artifact maps, all these are shown in Figure 3. To achieve even more realistic simulations we also added 30% of unit Gaussian noise to the subject's fMRI data.

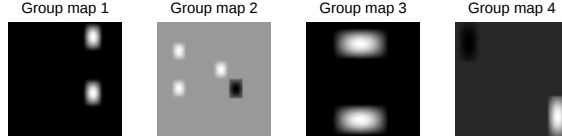


Figure 2: Representative fMRI maps of a group of subjects. Map 1 is task-related, maps 2 and 4 are transiently task-related and map 3 is functional related.

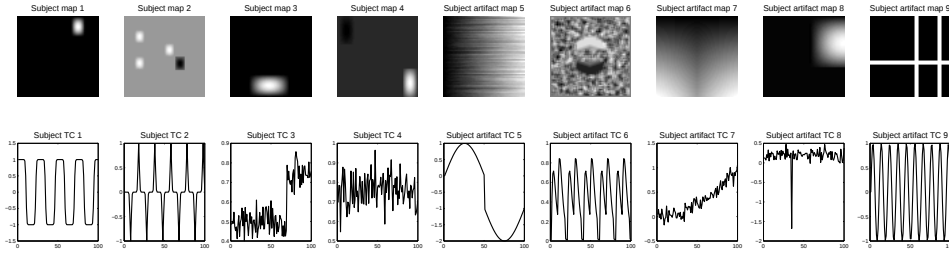


Figure 3: Shows a set of subject specific maps and their associated time-courses. This subject shares maps 1-4 with the group, however there are notable differences between the subject and the group in maps 1 and 3 where in each case one of the bright spots is missing. The other maps simulate artifacts including sinusoidal rotation motion (map 5), cardiac pulsation (map 6), scanner drift (map 7), head motion (map 8) and a strong scanner hardware artifact (map 9).

4.2. Stage 1 results

By fixing the 4 group maps the ICA outputs 11 spatial maps and their corresponding time-courses, as seen in Figure 4. Maps 1-4 are identical to the group maps, since these are the fixed group maps. The remaining maps are the estimated artifacts (maps 5-9) and the map differences between the subject and the group maps (maps 10 and 11).

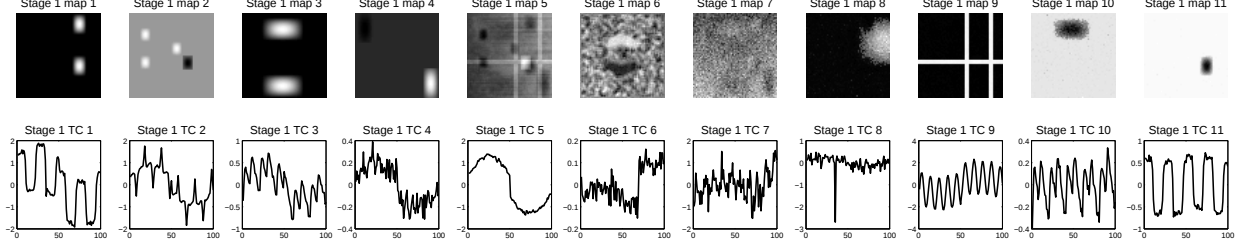


Figure 4: Shows the output of Dual-ICA after stage 1.

From the post-processing stage described in section 3.3, time-courses of maps 10 and 11 are identified as having a very high correlation to time-courses 3 and 1 respectively and as a result time-courses 8 and 10 are removed for stage 2.

4.3. Stage 2 results

After fixing the previously estimated time-courses, the final results from the Dual-ICA algorithm are shown in Figure 5. All four subject maps are found correctly with some added noise. The prominent artifacts (maps 6, 8 and 9 in Figure 5) are also estimated well.

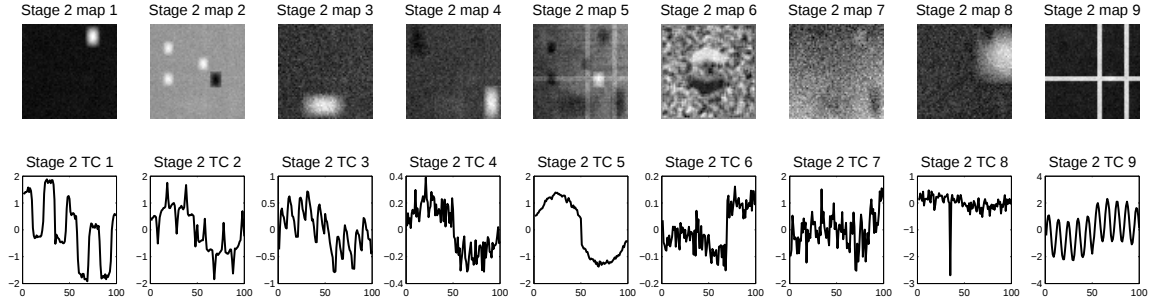


Figure 5: Shows the output of Dual-ICA after stage 2.

4.4. Comparison against dual regression

Dual regression is the current state of the art method for finding subject specific maps. It uses the group maps as a set of spatial regressors in a General Linear Model (GLM) which in this case is just a multivariate least squares regression, to find temporal dynamics associated with each group map. It then applies these time-courses as a set of temporal regressors in a GLM, to find subject specific maps [7]. The results from running dual regression on the simulated data yield heavily contaminated subject maps (bottom of Figure 6). This is because there is significant overlap and correlation between subject maps and time-courses, and artifact maps and time-courses, which causes the regression to pull some of the artifacts into the subject map estimates. The advantage of Dual-ICA is that it is able to separately estimate the extra artifact maps, giving them less chance to contaminate the subject maps (middle of Figure 6).

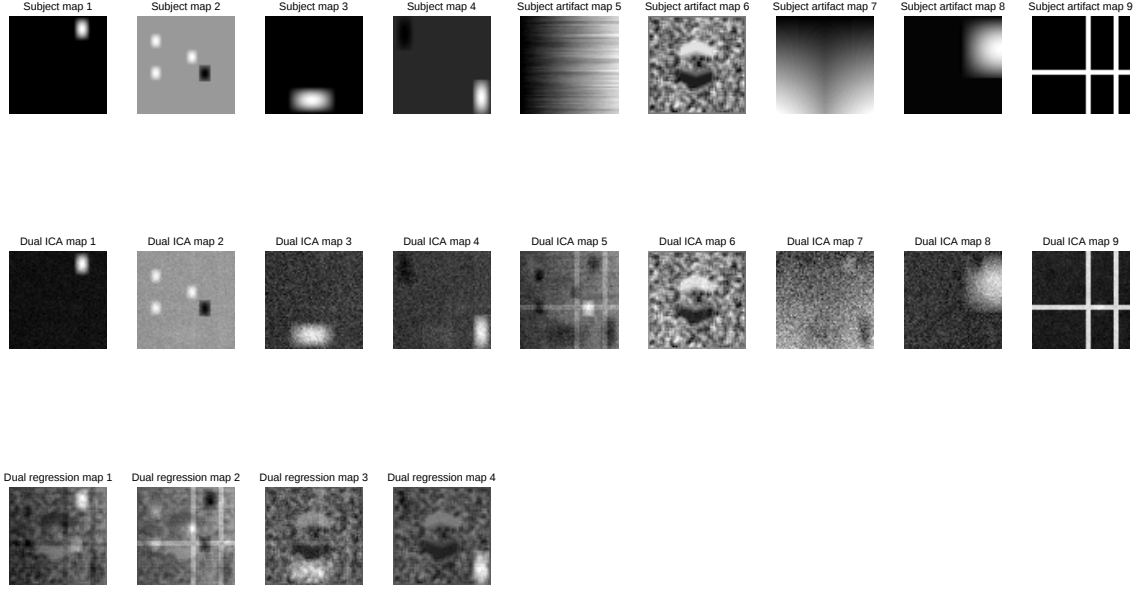


Figure 6: Shows the ground truth of the subject maps (top), maps found by Dual-ICA (middle) and maps estimated by dual regression (bottom).

Dual-ICA also performs well for different noise levels (Figure 7) and with varying number of artifacts added to the simulated subject data (Figure 8).

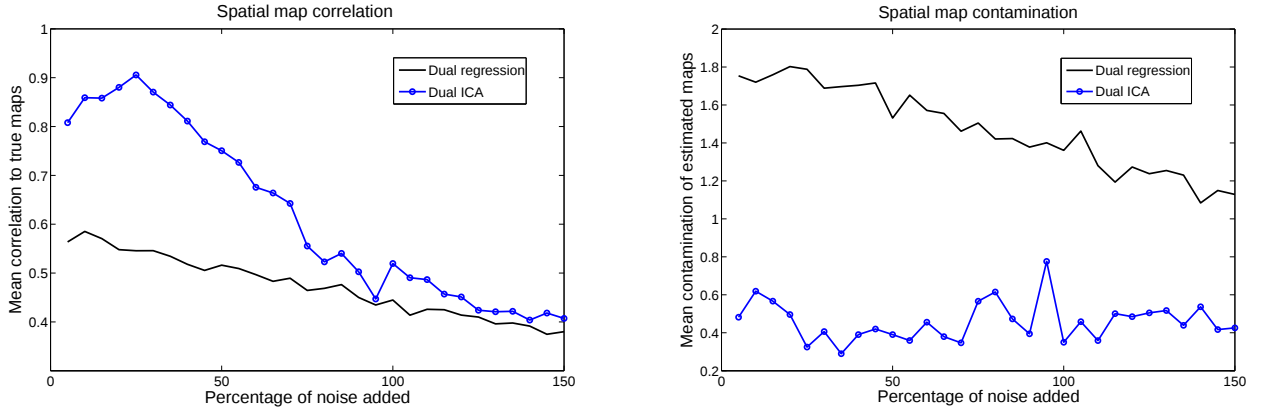


Figure 7: Shows the average amount of correlation between the estimated subject maps and the ground truth (left), and the total amount of contamination in the estimated subject maps (right) with varying amounts of noise added to the subject data. Here 100% corresponds to noise which is Gaussian and of the same (100%) magnitude as the spatial map.

To calculate how much a subject map is contaminated by a particular artifact, we first orthogonalise the artifact with respect to the subject map so that the artifact does not contain any subject map, and then calculate how much correlation there is between the orthogonalised artifact and the subject map. For all the correlations we used the well known Pearson correlation coefficient.

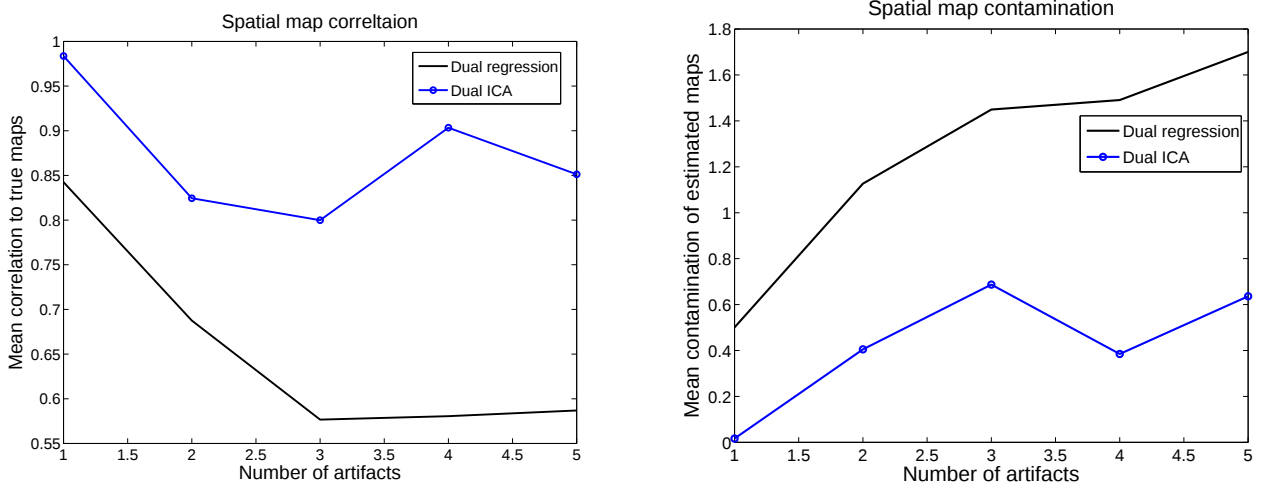


Figure 8: Shows the average amount of correlation between the estimated subject maps and the ground truth (left), and the total amount of contamination in the estimated subject maps (right) with varying number of artifacts in the subject data.

4.5. Real fMRI data

The group data used for this experiment was generated from 36 healthy participants. The subject specific fMRI data is taken from a participant that had some cognitive impairment. The Dual-ICA algorithm is able to consistently reproduce all the 20 subject specific network maps that are also present in the group maps. A sample output of one of the maps is shown in Figure 9. The proposed method also finds 4 additional artifacts in the data one of which is shown in Figure 10. Dual regression only regresses out subject maps that correspond to group maps and does not try to detect these additional artifacts.

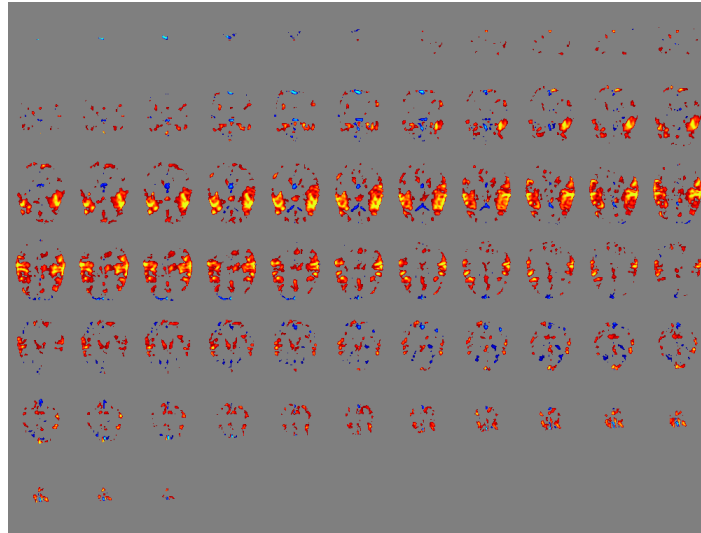


Figure 9: Shows one of the maps produced by the Dual-ICA algorithm, the different maps correspond to different slices of the brain. This particular map corresponds to the auditory system.

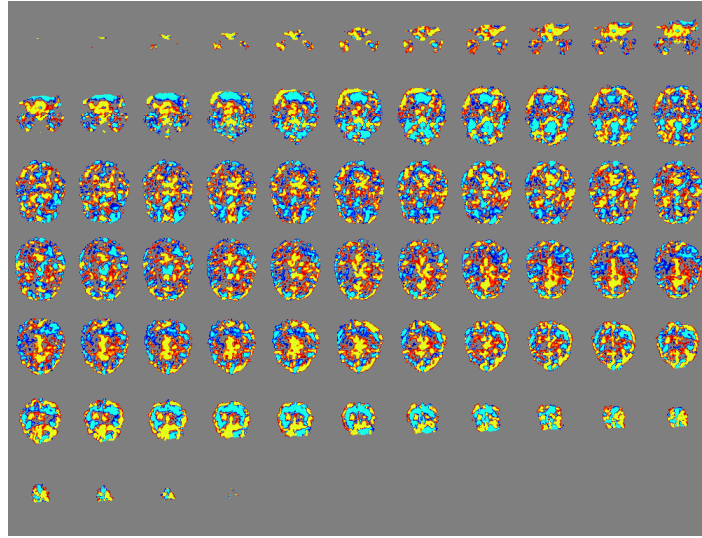


Figure 10: Shows one of the artifact maps found by Dual-ICA.

5. Future work

Through future work we will aim to properly evaluate these two methods on real fMRI by comparing results from larger studies and by using various statistical comparisons to test the accuracy of the maps, however this will be much more difficult for real fMRI than for simulated data because we don't actually know what the exact ground truth is.

The main disadvantages of this Bayesian Dual-ICA approach are the computational cost involved and the complexity of the implementation. Dual regression is extremely simple to implement and also fast at finding subject components making it accessible to many researchers. To make Bayesian Dual-ICA more practical, research into ways of simplifying the model and making good initialisations in order to decrease the number of iterations will be important.

6. Conclusions

Investigating functional brain networks (maps) has recently become a major research area within the neuroscience community. Dual regression has been used to decompose subject fMRI data into a set of time-courses and associated spatial maps by using existing group maps as regressors. However dual regression can often suffer from contamination due to various artifacts that are present in the subject's fMRI data.

In this work we presented a Bayesian Dual-ICA model to solve this contamination problem. By working within a Bayesian framework we are able to specify the priors in such way that it allows us to incorporate the group map information in order to guide the subject specific ICA to find independent spatial maps as well as artifact maps, thus preventing artifacts from contaminating the subject maps. The model was implemented and solved using Variational Bayes and tested on simulated and real fMRI data. We also compared our method against dual regression and for simulated data showed that it produces cleaner subject maps that have less spatial contamination from artifacts.

References

- [1] P. Jezzard, P. Matthews, S. Smith, et al., Functional MRI: an introduction to methods, volume 390, Oxford University Press Oxford, 2001.

- [2] M. McKeown, T. Sejnowski, et al., Independent component analysis of fmri data: examining the assumptions, *Human Brain Mapping* 6 (1998) 368–372.
- [3] A. Groves, C. Beckmann, S. Smith, M. Woolrich, Linked independent component analysis for multimodal data fusion, *Neuroimage* (2010).
- [4] A. Hyvärinen, E. Oja, A fast fixed-point algorithm for independent component analysis, *Neural computation* 9 (1997) 1483–1492.
- [5] R. Choudrey, S. Roberts, Flexible bayesian independent component analysis for blind source separation, in: *Proc. Int. Conf. on Independent Component Analysis and Signal Separation (ICA2001)*, Citeseer, pp. 90–95.
- [6] C. Beckmann, M. DeLuca, J. Devlin, S. Smith, Investigations into resting-state connectivity using independent component analysis, *Philosophical Transactions of the Royal Society B: Biological Sciences* 360 (2005) 1001.
- [7] C. Beckmann, C. Mackay, N. Filippini, S. Smith, Group comparison of resting-state fmri data using multi-subject ica and dual regression, in: *Proceedings of the 15th Annual Meeting of Organization for Human Brain Mapping*.
- [8] N. Filippini, B. MacIntosh, M. Hough, G. Goodwin, G. Frisoni, S. Smith, P. Matthews, C. Beckmann, C. Mackay, Distinct patterns of brain activity in young carriers of the apoe-ε4 allele, *Proceedings of the National Academy of Sciences* 106 (2009) 7209.
- [9] D. Wipf, S. Nagarajan, A new view of automatic relevance determination, *Advances in neural information processing systems* 20 (2008) 1625–1632.
- [10] C. Bishop, S. S. en ligne), *Pattern recognition and machine learning*, volume 4, Springer New York, 2006.
- [11] N. Lawrence, C. Bishop, Variational bayesian independent component analysis, Univ of Cambridge Tech Report (2000).

Appendix A. Initialisation of ICA

To start with we need to provide the iterative variational algorithm with some initialisation for the maps and time-courses. As with most ICA implementations, Group-ICA is initialised from a PCA decomposition. In stage 1 we use the group spatial maps as a set of spatial regressors, to find the time-courses associated with each group map

$$\mathbf{H} = \mathbf{Y}/\mathbf{G}. \quad (\text{A.1})$$

This is done using a simple multivariate least squares method. Next we subtract the group maps found within the subject’s fMRI data from this data and use PCA (implemented using a singular value decomposition) to find the most prominent time-courses $\hat{\mathbf{H}}$ left in the residuals

$$\hat{\mathbf{X}}\hat{\mathbf{\Sigma}}\hat{\mathbf{H}} = \text{SVD}(\mathbf{Y} - \mathbf{G}\mathbf{H}) \quad (\text{A.2})$$

Finally we use all the found time-courses $[\mathbf{H} \quad \hat{\mathbf{H}}]$ (subject and artifact) to regress out the corresponding spatial maps $[\mathbf{X} \quad \hat{\mathbf{X}}]$ (subject and artifact) from the data

$$[\mathbf{X} \quad \hat{\mathbf{X}}] = \mathbf{Y} / \begin{bmatrix} \mathbf{H} \\ \hat{\mathbf{H}} \end{bmatrix}. \quad (\text{A.3})$$

These spatial maps and time-courses are then used to initialise the iterative VB algorithm. A good initialisation is important so that the algorithm does not get stuck in a local minimum early on. For stage 2, where the time-courses are fixed a similar thing is done and the roles of the time-courses and spatial maps in the above procedure are just reversed.