

UNSUPERVISED DETRENDING TECHNIQUE USING SPARSE DICTIONARY LEARNING FOR fMRI PREPROCESSING AND ANALYSIS

Muhammad Usman Khalid

NICTA & The Australian National University
ANU College of Engg. & Computer Sci.
Canberra, Australia

muhammad.khalid@nicta.com.au

Abd-Krim Seghouane

Melbourne School of Engineering
Dept. of Electrical & Electronic Engg.
Melbourne, Australia

abd-krim.seghouane@unimelb.edu.au

ABSTRACT

This paper addresses the problem of scanner induced low frequency drift estimation in order to improve the significance of functional magnetic resonance imaging (fMRI) data for statistical analysis. A novel technique is presented to estimate the drift parameters using a sparse general linear model (sGLM) framework. The fMRI signal is modeled as a linear mixture of several signals such as low frequency trend, brain hemodynamic, physiological noise and unexplained signal variations. These signals are considered as underlying sources and sparse dictionary learning (SDL) is used to estimate them. The superior performance of the proposed technique compared to other detrending techniques is illustrated using a simulation study. Furthermore, the proposed technique is validated using real fMRI data, which shows its better capability to estimate drift in presence of spatiotemporal dependencies.

Index Terms— K-SVD, CCA, DCT, fMRI, detrending

1. INTRODUCTION

The tasks associated with fMRI statistical analysis generally consists in answering three questions i) which areas of the brain are activated in response to a given stimulus [1], ii) what is it the temporal dynamics of the activated brain areas during activation [2] and iii) how are the connections between the different activated brain areas [3]. For fMRI, GLM is a widely accepted hypothesis-driven mass-univariate approach to analyze regional brain activity [4]. It requires regressors for its design matrix that includes canonical hemodynamic response function (HRF) and its derivatives convolved with a stimulus function [5]. Due to lack of prior knowledge about HRF variability, over subjects and experiments [6], and other unexplained variations, data-driven methods are considered to provide an alternative to hypothesis based fMRI analysis. The data-driven exploratory techniques such as principal component analysis (PCA) [7], independent component analysis (ICA) [8] and canonical correlation analysis (CCA) [9] aim at exploring unique hidden patterns based on covariations

in the multivariate data such that task-related blood-oxygen-level dependent (BOLD) time-series and other time-varying effects can be separated. Nevertheless, they have shown inferior results to sparsity based learning that better captures the spatiotemporal characteristics [10, 11, 12, 13, 14].

However, the presence of lag one autocorrelations due to low frequency drifts can compromise the performance of sparsity based learning. Their elimination from fMRI data is substantial for better results from SDL. If not removed, their non-sparse nature will cause trained dictionary atoms to contain lag one autocorrelations. Previous studies have shown that drifts exhibit high global correlations [15], thus causing failure of SDL algorithms in disintegration of temporal mixtures of brain hemodynamics from non-hemodynamics on the basis of a sole assumption of sparsity [16]. To resolve this problem sGLM model assumed a unique drift at each voxel and used a linear combination of discrete cosine transform (DCT) basis functions to prefilter trends from the data [10].

On the other hand CCA based exploratory technique assumes a global drift and retrieves it on the basis of autocorrelation maximization which can be pre-filtered from the fMRI data [17]. Using a similar approach, we allow SDL itself to estimate the drift source. This retrieved drift can be pre-filtered from the fMRI data using least square fit. The proposed technique has been built upon the existing framework of sGLM and therefore can be used for task related activation detection and resting state functional connectivity analysis.

2. METHODS

For the sparse GLM model, consider a fMRI BOLD signal \mathbf{x}_i at the i -th voxel over the course of N scanned volumes to be represented as a linear combination of exactly or less than k atoms from an unknown dictionary $\mathbf{A} \in \mathbb{R}^{N \times l}$ on the basis of sparse response signal strength $\phi_i \in \mathbb{R}^l$,

$$\mathbf{x}_i = \mathbf{A}\phi_i + \mathbf{e}_i, \quad i = 1, 2, \dots, v \quad (1)$$

where $N \ll l$, the residual $\mathbf{e}_i \in \mathbb{R}^N$ is a random Gaussian noise vector with variance σ_i^2 , distributed as $\mathcal{N}(0, \sigma_i^2 \mathbf{A})$,

with temporal correlation \mathbf{A} that accounts for the correlated noise. Here, drift is modeled as a realization of long memory noise process and different detrending techniques are used for its elimination from the data. A Gaussian filter based pre-coloring is used for the removal of high frequency auto-correlations. The correlation matrix \mathbf{A} can then be assumed as an identity matrix $\mathbf{I}_N \in \mathbb{R}^{N \times N}$. After detrending and pre-coloring, the model in (1) can be addressed using SDL algorithm such as [11] to learn \mathbf{A} and Φ .

Algorithm 1: Identifying drift among CCA components

Given: $\mathbf{X}_d \in \mathbb{R}^{N \times v}$,

1. Compute DCT basis [18] to obtain drift matrix \mathbf{A}_t using $\mathbf{b}_m = \sqrt{\frac{2}{N}} \text{Cos}(\frac{\pi(2n+1)m}{2N})$, where $m = 1, \dots, M$, $n = 0, 1, \dots, N-1$, $M = \lfloor \frac{2N}{f_s f_c} + 1 \rfloor$ is the number of basis, f_s is the sampling frequency, and f_c is the cut-off frequency, respectively

$$\mathbf{A}_t = \mathbf{B}\mathbf{B}^T \mathbf{X}_d \quad (2)$$

2. For a normalized \mathbf{X}_d , solve $\mathbf{X}_d^T \mathbf{X}_d \mathbf{u} = \lambda \mathbf{u}$, where λ and \mathbf{u} are the eigenvalues and eigenvectors, respectively, to obtain principal components along temporal dimension [19] as, $\mathbf{X}_p = \mathbf{X}_d \mathbf{u}$
3. Use canonical correlation [19] to measure autocorrelation between \mathbf{X}_p and $\mathbf{Y}_p = \mathbf{X}_p(n-1)$, by solving $\Sigma_{X_p X_p}^{-1} \Sigma_{X_p Y_p} \Sigma_{Y_p Y_p}^{-1} \Sigma_{Y_p X_p} \mathbf{u} = \rho \mathbf{u}$, where ρ and \mathbf{u} are canonical coefficients and weights, respectively, and Σ signifies a correlation matrix, to obtain, $\mathbf{X}_c = \mathbf{X}_p \mathbf{u}$
4. Regression analysis between \mathbf{X}_c and \mathbf{A}_t ,

$$\mathbf{a}_{t_i} = \mathbf{X}_c \boldsymbol{\omega}_i + \boldsymbol{\epsilon}, \quad i = 1, 2, \dots, v, \quad (3)$$

5. Calculate $\max(\text{diag}(\boldsymbol{\Omega} \boldsymbol{\Omega}^T))$ to identify location of drift \mathbf{a}_c among CCA components.
-

The proposed technique for detrending is applied on pre-colored and nondetrended data-set \mathbf{X}_d followed by further analysis for statistical inferences. This whole scheme consists of four steps i) learning drift from the training data during the first pass of SDL under the assumption of an overcomplete dictionary, $N \ll l$, ii) locating drift among dictionary atoms using the combination of DCT and CCA, so that the whole scheme can be executed in an unsupervised manner, iii) removing drift from the training data using least square fit, and iv) relearning dictionary from the detrended data. During the first pass, a dictionary \mathbf{A}_d is learned from data-set \mathbf{X}_d , which contains two kinds of temporal profiles labeled as spatially non-integrated components \mathbf{A}_{nc} and spatially integrated components \mathbf{A}_{ic} . The components from \mathbf{A}_{nc} correspond to the subspace pertaining to underlying sources including resting state activities, low frequency drifts and other interfering nuisance components. On the other hand, \mathbf{A}_{ic} contains temporal mixtures of underlying sources, which fall under the cate-

gory of signal integration [20] and within these components task related hemodynamics are mixed with other signal variations such as drifts. Our goal is to somehow detect the non-integrated dictionary atom \mathbf{a}_d that consists solely of drift, so that it can be used to detrend the entire data-set. Once it is separated from $\mathbf{A}_{nc} \in \mathbb{R}^{N \times l}$, we are left with $\mathbf{A}_{nc}^* \in \mathbb{R}^{N \times l-1}$. In order to identify the dictionary atom that consists of a low frequency drift, consider the following model

$$\begin{aligned} \mathbf{X}_d &= \mathbf{A}_d \boldsymbol{\Psi}_d + \mathbf{E} \\ &= \mathbf{A}_{nc} \boldsymbol{\Psi}_{nc} + \mathbf{A}_{ic} \boldsymbol{\Psi}_{ic} + \mathbf{E} \\ &= \mathbf{a}_d \mathbf{z}_d + \mathbf{A}_{nc}^* \boldsymbol{\Psi}_{nc}^* + \mathbf{A}_{ic} \boldsymbol{\Psi}_{ic} + \mathbf{E} \\ &= \mathbf{a}_d \mathbf{z}_d + \mathbf{A} \Phi + \mathbf{E} \\ &= \mathbf{a}_d \mathbf{z}_d + \mathbf{X} \end{aligned} \quad (4)$$

Algorithm 2: Learning and identifying drift atom, detrending fMRI data-set and inferences

Given: $\mathbf{X}_d \in \mathbb{R}^{N \times v}$, k , and \mathbf{a}_c

1. Estimate \mathbf{A}_d and $\boldsymbol{\Psi}_d$ from (4) using SDL,
2. Perform correlation analysis between \mathbf{a}_c and \mathbf{A}_d to identify the drift \mathbf{a}_d among dictionary atoms, σ and γ signifies standard deviations and correlation values, respectively, $\gamma = \max \left| \frac{\mathbf{a}_c^T \mathbf{A}_d}{\sigma_{\mathbf{a}_c} \sigma_{\mathbf{A}_d}} \right|$
3. From (6) find \mathbf{z}_l and use (7) to obtain \mathbf{X} ,
4. Run the second pass of SDL using model (1),
5. Use F-test for inferences, where \mathbf{A}_{s_i} is a local dictionary, $\mathbf{P}_{\mathbf{A}_{s_i}}^\perp$ is the projection matrix associated with the subspace of \mathbf{A}_{s_i} , $\mathbf{A}_{s_i \setminus \mathbf{z}}$ is a reduced size design matrix of a local dictionary, M is the rank of \mathbf{A} , and $\{\mathbf{s}_i\}_{i=1}^v$ captures the indices of k most correlated dictionary atoms with the data \mathbf{X} [10],

$$\mathbf{F}_i = \frac{\mathbf{x}_i^T (\mathbf{P}_{\mathbf{A}_{s_i \setminus \mathbf{z}}}^\perp - \mathbf{P}_{\mathbf{A}_{s_i}}^\perp) \mathbf{x}_i}{\mathbf{x}_i^T \mathbf{P}_{\mathbf{A}_{s_i}}^\perp \mathbf{x}_i} (N - M). \quad (5)$$

where $\mathbf{z}_d \in \mathbb{R}^v$ is the corresponding sparse strength of drift atom, $\mathbf{A} = \mathbf{A}_{nc}^* + \mathbf{A}_{ic}$, $\Phi = \boldsymbol{\Psi}_{nc}^* + \boldsymbol{\Psi}_{ic}$ and $\mathbf{X} = \mathbf{A} \Phi + \mathbf{E}$. The model in (4) can be solved using $\min_{\phi_i} \|\mathbf{x}_{d_i} - \mathbf{A} \psi_{d_i}\|_2^2$, subject to $\|x_i\|_0 \leq k$. Finding the optimal k corresponds to a problem of model selection criterion that can be resolved using a univariate model selection criterion [21, 22, 23, 24]. We considered a K-SVD algorithm to address the model (4) to learn drift subspace \mathbf{a}_d and remaining signal subspace \mathbf{A}_{nc}^* and \mathbf{A}_{ic} . The detection of drift atom in \mathbf{A}_d is not a straight forward procedure, because unlike exploratory data-driven techniques that arrange the transformed components according to some criteria for instance PCA orders them by energy and CCA by autocorrelation, dictionary atoms are not arranged by any specific order. Therefore, DCT basis along with a CCA based drift estimator are deployed simultaneously to allow an unsupervised execution of the scheme. As CCA orders its components according to maximum autocorrelation, therefore drift will always be

the first component. However, when a blocked experimental design is employed we may get into a risk of estimating BOLD response as first component due to more probability of its higher autocorrelation. This problem was addressed by using linear regression analysis between DCT basis and CCA, which provides the exact location of drift among CCA components. Using DCT basis directly to identify drift atom is not advisable because \mathbf{A}_d contains drift not only as an independent atom \mathbf{a}_d , but it may also be linearly mixed into other sources contained in sub-dictionary \mathbf{A}_{ic} . Furthermore, only

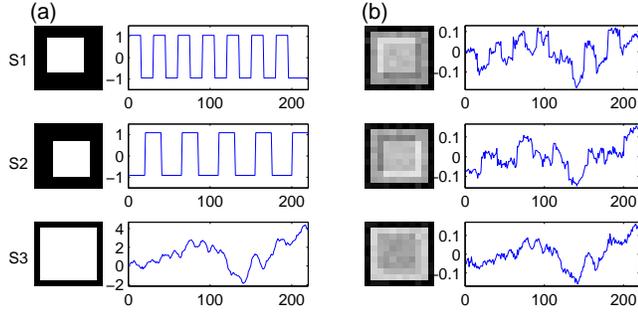


Fig. 1. a) Sources, b) Retrieved sources by SDL.

retaining the subspace \mathbf{a}_d requires re-estimation of its corresponding signal strength \mathbf{z}_d . Alternatively, signal strength \mathbf{z}_l is estimated using least squares.

$$\min_{\mathbf{z}} \|\mathbf{a}_d \mathbf{z}_l - \mathbf{x}_{d_i}\|_2^2 \quad (6)$$

If \mathbf{X}_c are the components obtained by CCA technique using $\mathbf{X}_c = \mathbf{X}_p \mathbf{u}$ and \mathbf{A}_t are the drift estimates obtained at each voxel using (2) then (3) allows us to estimate strength of relationship Ω between \mathbf{x}_{c_j} and \mathbf{a}_{t_i} , where $j=1,2,\dots,J$, $i=1,2,\dots,v$, J being the total number of CCA components. The auto-covariance matrix of Ω provides a quantifiable measure that can be used to locate the drift \mathbf{a}_c among CCA components. The full procedure for second step of the scheme is described in Algorithm 1. Due to the curse of dimensionality, the number of voxels being much greater than the number of observations, PCA has been used to prewhiten the data by keeping only few components with highest variance.

In the third step, temporal correlation analysis between drift component of CCA \mathbf{a}_c and dictionary atoms \mathbf{A}_d allows the detection of atom that contains drift. The dictionary atom in \mathbf{A}_d that maximally correlates with \mathbf{a}_c is considered as the drift atom \mathbf{a}_d . It is used to remove trends from the data using least square projection given as

$$\mathbf{X} = \mathbf{X}_d - \mathbf{a}_d \mathbf{z}_l \quad (7)$$

where $\mathbf{X} \in \mathbb{R}^{N \times v}$ is the detrended fMRI data, which is considered for sparse GLM modeling to perform statistical inferences such as activation detection. The details about third and fourth step of the proposed scheme followed by an F-test for statistical inferences are given in Algorithm 2.

3. APPLICATION

3.0.1. Simulation Study

The simulation study allowed us to examine and compare the strength of DCT, exploratory and SDL techniques for drift estimation in presence of spatiotemporal dependencies. Three waveforms each consisting of 220 time-points were assigned to three different slices according to activation patterns $S1 - S3$. The patterns $S1 - S3$ consisted of patches 11×11 pixels

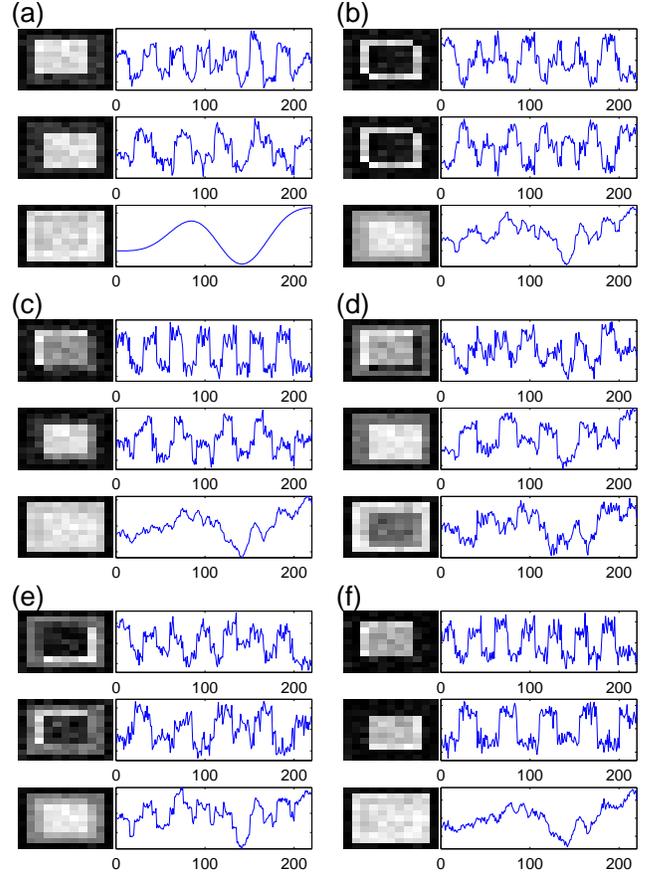


Fig. 2. Retrieved sources $S1 - S2$ by SDL (when applied to six different detrended data-sets) and estimated drifts by a) DCT, b) PCA, c) CCA, d) sICA, e) tICA, f) proposed shown in top two rows and bottom row of each subfigure, respectively. The drift for DCT is obtained by taking SVD of drift estimates.

with an activation of amplitude 1 from pixel 3–8 for $S1$, 4–9 for $S2$ and 2–11 for $S3$ along both dimensions, respectively, as shown in Fig 1a. According to activation sources, the source waveforms along with random white gaussian noise $\mathcal{N}(0, 0.3)$ were added together to generate a mixture of time-series called the generated data-set. The waveform $S3$ is modeled as drift and generated using long memory noise.

The first step consisted of source estimation using SDL whose results are shown in Fig. 1b. The drift source $S3$

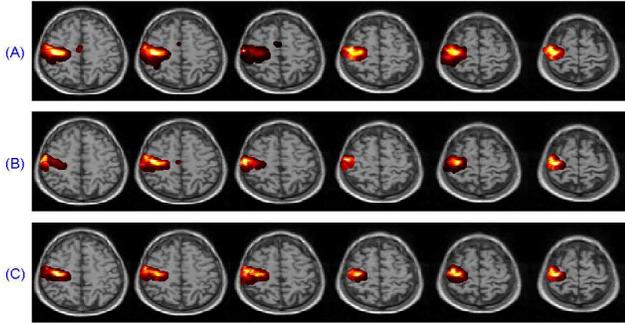


Fig. 3. Activation maps for block design RFT task at a p -value threshold of 0.0001 using sGLM framework for detrended data-set by (A) DCT, (B) CCA, and (c) proposed.

due to its non-sparseness could not be decorrelated from other sources by SDL. However, it was correctly estimated and used to detrend the generated data-set. In the second step drift was retrieved using DCT, PCA, CCA, sICA, and tICA followed by detrending of generated data-set using their corresponding retrieved drifts. In the third step SDL was reapplied on all detrended data-sets. The results from third step consisted of retrieved activation patterns and their corresponding waveforms S_1 and S_2 , which are illustrated in Fig. 2. For all techniques, the accuracy of extracted sources S_1 and S_2 by SDL demonstrates how precisely each technique was able to estimate S_3 . The spatiotemporal preciseness of retrieved sources by SDL showed its superior drift estimation capability over other multivariate techniques.

3.0.2. Real fMRI data

The K-SVD algorithm was chosen for dictionary learning due to its superior performance, and thresholding correlation [25] was used for sparse coding due to its computational simplicity. For an unbiased comparison, fixed number of iterations were used for the convergence of K-SVD. The real data-sets consisted of A) block-related right finger tapping (RFT) task dataset: TR = 3 sec, with a total acquisition time of 480 sec, and B) event-related RFT task dataset: TR = 2 sec, with a total acquisition time of 650 sec. The acquisition details about these data-sets can be found in [10]. The image pre-processing for real fMRI data was carried out in Matlab that consisted of five steps, i) realignment, ii) normalization, iii) spatial smoothing, iv) masking, and v) temporal smoothing. All functional images were realigned to the first image in order to correct for any head movements that may have occurred during the course of the experiment. In the next step, all images were spatially normalized to a standard Talairach template, resampled to $2mm \times 2mm \times 2mm$ voxels and spatially smoothed using a $8mm \times 8mm \times 8mm$ full-width at half-maximum (FWHM) Gaussian kernel. To remove any data outside the scalp images were masked and only those voxels were kept which exceeded a masking threshold.

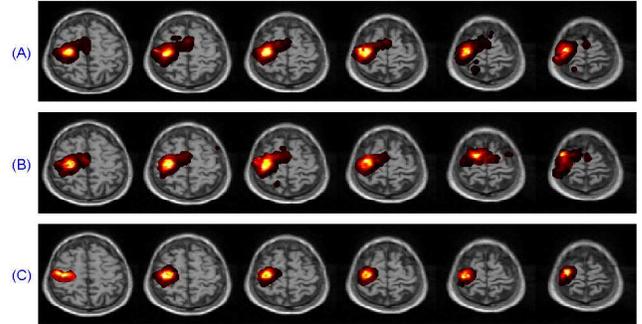


Fig. 4. Activation maps for event design RFT task at a p -value threshold of 0.0001 using sGLM framework for detrended data-set by (A) DCT, (B) CCA, and (c) proposed.

The 4-dimensional data-sets collected from the masked images were re-arranged as 3-dimensional matrices, and stored in $\mathbf{X}_d \in \mathbb{R}^{N \times v \times S}$ to be used as a whole brain's measured data arranged according to the slice numbers $s = 1, 2, \dots, S$. The autocorrelations in the data due to high frequencies were removed by temporal smoothing, where 1.5 s FWHM Gaussian filter was used. The data-sets from each slice were then used for slice based analysis. Considering an unbiased comparison for SDL, a total of 30 dictionary atoms were trained, using 30 number of iterations for each slice. For exploratory techniques such as CCA first 30 components from PCA were used.

The detrending by DCT, CCA, and proposed was followed by SDL. For SDL, according to AIC criteria sparsity parameter k for block and event design was set to 2, and 4, respectively. For activation detection, the most correlated atom with the modeled BOLD hemodynamic response (MHR) along with other non-zero $k-1$ dictionary atoms were chosen as regressors for sparse GLM. The MHR is obtained by convolving a canonical HRF with a stimulus function. The results illustrated in Fig. 3 and 4 reveal that the neural activations in motor area for RFT are truly identified for all techniques with few exceptions. The results of activation detection based on proposed detrending showed more specific activation maps without false activations.

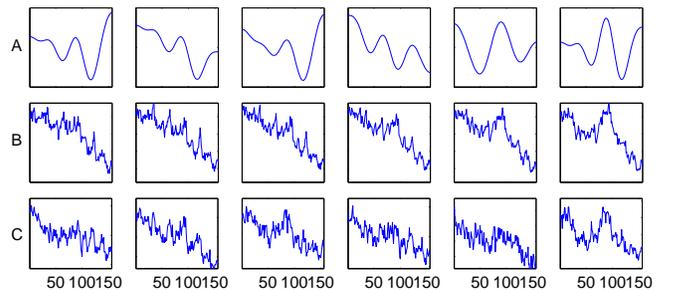


Fig. 5. Estimated drifts for block design RFT task using (A) DCT, (B) CCA, and (c) proposed. The drift for DCT is obtained by taking SVD of drift estimates at relevant pixels.

4. REFERENCES

- [1] A. K. Seghouane and J. L. Ong, "A Bayesian model selection approach to fMRI activation detection," *IEEE International Conference on Image Processing*, pp. 4401–4404, 2010.
- [2] A. K. Seghouane and A. Shah, "HRF estimation in fMRI data with unknown drift matrix by iterative minimization of the Kullback-Leibler divergence," *IEEE Transactions on Medical Imaging*, vol. 31, pp. 192–206, 2012.
- [3] A. K. Seghouane and S. I. Amari, "Identification of directed influence: Granger causality, Kullback-Leibler divergence and complexity," *Neural Computation*, vol. 24, pp. 1722–1739, 2012.
- [4] K. J. Friston, J. Ashburner, S. Kiebel, T. Nichols, and W. Penny, *Statistical Parametric Mapping: The Analysis of Functional Brain Images*, Academic Press, 2006.
- [5] B. Lenoski, L. C. Baxter, L. J. Karam, J. Maisog, and J. Debbins, "On the performance of autocorrelation estimation algorithms for fMRI analysis," *IEEE Journal of sel. Top. in sig. proc.*, vol. 2, pp. 828–838, 2008.
- [6] G. K. Aguirre, E. Zarahn, and M. D'Esposito, "The variability of human BOLD hemodynamic responses," *NeuroImage*, vol. 8, pp. 360–369, 1998.
- [7] K. J. Friston, C. D. Frith, P. F. Liddle, and R. S. Frackowiak, "Functional connectivity: the principal component analysis of large PET data sets," *Journal of Cerebral Blood Flow and Metabolism*, vol. 13, pp. 5–14, jan 1993.
- [8] M. J. Mckeown, S. Makeig, G. G. Brown, T. P. Jung, S. S. Kindermann, A. J. Bell, and T. J. Sejnowski, "Analysis of fMRI data by blind separation into independent spatial components," *Human Brain Mapping*, vol. 6, pp. 160–188, 1998.
- [9] O. Friman, M. Borga, P. Lundberg, and H. Knutsson, "Exploratory fmri analysis by autocorrelation maximization," *Neuroimage*, vol. 16, pp. 454–464, 2002.
- [10] K. Lee, S. Tak, and J. C. Ye, "A data-driven sparse GLM for fMRI analysis using sparse dictionary learning with MDL criterion," *IEEE Trans. Med. Imaging*, vol. 5, pp. 1076–1089, 2010.
- [11] M. U. Khalid and A. K. Seghouane, "Constrained maximum likelihood based efficient dictionary learning for fMRI analysis," *International Symposium on Biomedical Imaging*, pp. 45–48, 2014.
- [12] M. U. Khalid and A. K. Seghouane, "A single SVD sparse dictionary learning algorithm for fMRI data analysis," *Statistical Signal Processing*, pp. 65–68, 2014.
- [13] M. Hanif and A. K. Seghouane, "Maximum likelihood orthogonal dictionary learning," *IEEE Workshop on Statistical Signal Processing (SSP)*, pp. 1–4, 2014.
- [14] M. U. Khalid and A. K. Seghouane, "Multiple-subject fMRI connectivity analysis using sparse dictionary learning and CCA," *In the Proceedings of International Symposium on Biomedical Imaging*, 2015.
- [15] L. Yan, Y. Zhuo, Y. Ye, S. X. Xie, J. An, G. K. Aguirre, and J. Wang, "Physiological origin of low-frequency drift in BOLD fMRI," *Magn. Reson. Med.*, vol. 61, pp. 819–827, 2009.
- [16] M. U. Khalid and A. K. Seghouane, "Improving functional connectivity detection in fMRI by combining sparse dictionary learning and canonical correlation analysis," *International Symposium on Biomedical Imaging*, pp. 286–289, 2013.
- [17] O. Friman, M. Borga, P. Lundberg, and H. Knutsson, "Detection and detrending in fMRI data analysis," *Neuroimage*, vol. 22, pp. 645–655, 2004.
- [18] N. Ahmed, T. Natarajan, and K. Rao, "Discrete cosine transform," *IEEE Trans. Comput.*, vol. 23, pp. 90–93, 1974.
- [19] F. G. Ashby, *Statistical analysis of fMRI data*, MIT Press, 2011.
- [20] A. Bruguier, K. Preuschoff, S. Quartz, and P. Bossaerts, "Investigating signal integration with canonical correlation analysis of fMRI brain activation data," *Neuroimage*, vol. 41, pp. 35–44, 2008.
- [21] A. K. Seghouane and M. Bekara, "A small sample model selection criterion based on the kullback symmetric divergence," *IEEE Transactions on Signal Processing*, vol. 52, pp. 3314–3323, 2004.
- [22] A. K. Seghouane and S. I. Amari, "The AIC criterion and symmetrizing the Kullback-Leibler divergence," *IEEE Transactions on Neural Networks*, vol. 18, pp. 97–106, 2007.
- [23] A. K. Seghouane and A. Cichocki, "Bayesian estimation of the number of principal components," *Signal Processing*, vol. 87, pp. 562–568, 2007.
- [24] A. K. Seghouane, "Asymptotic bootstrap corrections of AIC for linear regression models," *Signal Processing*, vol. 90, pp. 217–224, 2010.
- [25] R. Gribonval, H. Rauhut, K. Schnass, and P. Vandergheynst, "Atoms of all channels, unite! average case analysis of multi-channel sparse recovery using greedy algorithms," *Journal of Fourier Analysis and Applications*, vol. 14, pp. 655687, 2008.