

# Analysis of Auditory fMRI Recordings via ICA: A Study on Consistency

Jarkko Ylipaavalniemi  
Neural Networks Research Centre  
Helsinki University of Technology  
P.O. Box 5400, FI-02015 HUT, Finland  
E-mail: [whyj@cis.hut.fi](mailto:whyj@cis.hut.fi)

Ricardo Vigário  
Neural Networks Research Centre and  
Advanced Magnetic Imaging Centre  
Helsinki University of Technology  
P.O. Box 5400, FI-02015 HUT, Finland  
E-mail: [ricardo.vigario@hut.fi](mailto:ricardo.vigario@hut.fi)

**Abstract**— We apply a blind source separation approach to the identification of statistically independent spatial patterns of brain activation to auditory stimulation. Stimuli consisted of spoken text. The data was collected via functional magnetic resonance imaging (fMRI).

As expected from standard processing of fMRI, we observe that independent component analysis (ICA) reveals spatial patterns with similar temporal activation as the stimulus. In these, ICA further distinguishes between the primary auditory areas and Broca’s and Wernicke’s, which are associated with speech production and understanding, respectively. Furthermore, we observe the activation of the thalamus, with a time course unrelated to the stimulus, hence hard to detect in a classical manner. We observe as well a temporally evolving artifact, related to inefficient filtering of the fMRI scans.

The consistency of the estimated signals is tested by running the algorithm with many different initial conditions. The solutions found are combined according to their similarities. Estimates that differ greatly from run to run are less likely to correspond to true components, whereas those that present small variances are considered reliable ones.

## I. INTRODUCTION

Functional magnetic resonance imaging (fMRI) provides an indirect and non-invasive measure of neuronal activity based on the blood oxygenation level dependent (BOLD) signal [1], which is a measure of local changes in blood flow and oxygen consumption. In addition to being related to neuronal activity in a complex and time delayed way, this volumetric signal also includes contributions from other sources. These include artifacts related to poor scanning and head movements, as well as from physiological phenomena unrelated to the study, such as the heart beat or breathing.

Traditionally, experiments have tested brain responses to carefully designed repetitions of stimulus and resting sequences. Usually, stimuli are in the form of auditory, somatosensory or visual patterns. The experiment may include as well a simple task, like naming a color or moving a finger. The data is then analyzed using hypothesis-driven techniques, relating the brain activation with a predicted reference signal, based on the temporal design of the stimulation (*c.f.* [2]).

Hence, the study of brain regions other than the primary cortical areas has been difficult, since the activity of non-primary and sub-cortical areas is more complex and not so easy to relate to the given stimulus pattern. In Sec. IV we give

an example of activation of secondary auditory areas, which have a temporal pattern only mildly related to the stimulus.

The advent of new signal processing techniques, such as blind source separation (BSS), has enabled the analysis of more complex and less predictable brain activations [3], [4], [5]. Among such data-driven methods, independent component analysis (ICA, [6]) plays an important role. This approach, summarized in Sec. II-A, has seen a growing number of successful applications, both in medical and other areas. However, ICA is an adaptive algorithm which optimizes a given contrast function. Due to a limited computing accuracy, it may produce slightly different estimates of the sought solutions, depending on the initialization of its parameters. In the present study we actively exploit this variability, and suggest a way to assess the consistency of the independent components.

In an auditory paradigm, where the subject listens to spoken text, we show that consistent components include the expected ones, with stimulus-related activation patterns originating from the primary cortical areas, but also others, whose time courses are only mildly related to the stimulus. Additionally, our approach reveals new phenomena not directly related to the stimulus.

## II. INDEPENDENT COMPONENT ANALYSIS

In general, BSS attempts to estimate underlying sources from linearly mixed observations. The term ‘blind’ refers to the fact that very little to no *a priori* information exist on both the sources or the mixing process. Under certain assumptions, the problem can be solved with temporal decorrelation methods, such as SOBI or TDSEP [7], [8]. ICA is another family of methods for BSS, based on the assumption of statistical independence of the source signals (for a good textbook on the matter, *c.f.* [6]).

### A. Spatial and Temporal ICA

Let the fMRI signal be represented by the data matrix  $\mathbf{X}_{tj}$ , where  $t = 1, \dots, T$  ( $T$  denoting the total number of collected time points) and  $j = 1, \dots, J$  ( $J$  denoting the total number of voxels). Here we assumed that each volume, at any given time, has been vectorized in an unique manner.

The linear mixing model can then be expressed in the form:

$$\mathbf{X}_{tj} = \sum_{k=1}^K \mathbf{A}_{tk} \mathbf{S}_{kj}, \quad (1)$$

where each row of  $\mathbf{S}$  contains a statistically independent volume, and the corresponding column of the mixing matrix  $\mathbf{A}$  holds its temporal activations. Here we make the statistical constraint in the spatial domain. In some other applications, independence applies to the temporal domain, which leads into a transposed version of model (1).

Note that, since both  $\mathbf{A}$  and  $\mathbf{S}$  are unknown, we are free to include the signs and scaling of the independent components in either one. Also, the order of the components is not determined.

It is not the purpose of this paper to focus on methods to solve the ICA problem. Yet, in order to understand the following, one can give an intuitive explanation to the fundamentals of ICA, based on the search for maximally non-gaussian source estimates. According to the central limit theorem, mixed signals are more gaussian than any of the original sources. Hence, maximizing the non-gaussianity of the estimates should lead into their independence. A typical measure of non-gaussianity, which will be used in the present study, is the kurtosis. A relation between this explanation and the more theoretically grounded mutual information can be found, *e.g.*, in [6].

In FastICA [9], [10], as in many other ICA algorithms, the initial step consists of whitening the data. This results in uncorrelated and unit-variance data, which facilitates the separation of the underlying independent source signals. Furthermore, it has been shown that a suitable reduction of dimensionality at this stage may help prevent undesirable overfitted solutions [11].

### B. Exploiting the variability in FastICA

In model (1) we have made a few assumptions which are not always guaranteed in practice. For example, the measurements are noise free and the source components statistically independent. The latter assumption has been studied in [12], [13]. If it fails, it can be replaced by a somewhat less restrictive constraint of the active brain areas to be spatially sparse, as suggested in [14]. Yet, noise may render the identification of the true underlying sources difficult. Furthermore, the algorithm may approach the solution via different paths on the error surface, depending on its initial conditions.

The magnitude and nature of component variability can be analyzed by running the ICA algorithm multiple times and making sure that the starting conditions differ in each run. This is particularly suitable for an algorithm such as FastICA, due to its deflationary nature and convergence speed. Furthermore, in relation to fMRI studies, it has been shown that choosing a particular algorithm implementation of ICA is not crucial [15], *i.e.* their results do not vary greatly.

In the ideal situation, where the same true solution is reached in more than one run, these estimates should have

a cross-correlation of  $\pm 1$ . If the correlation is not perfect, but still high, and the number of runs is sufficiently big, one may expect that the true solution exists in an average of all these estimates. If, on the other hand, there is a big spread of correlations, then we are probably in the presence of not so consistent components. Let us consider the extreme case of completely overfitted estimates. In ICA, these correspond to a bump or a spike, which can occur anywhere in the data [11]. It is then very likely that they will occur in a different location in each run, hence the components will present a very poor correlation set.

We can then say that the number of times a particular component has been estimated, plus the spread of correlations of similar components give important information, which can be used for ranking the reliability of a given component.

Note that, using the averaging method, we are allowed as well to depart somewhat from the strict assumption of spatially independent components. Furthermore, because of the deflationary nature of FastICA, the estimate of the  $n$ th independent component is subject to the accumulated errors of estimating all  $n - 1$  earlier components. Thus, in the ultimate case of running FastICA to find only a single component at a time, this influence is minimized, as all components found are the very first ones.

### C. Implementation issues

Running ICA many times for the same data is straightforward. One only needs to ensure that the initial conditions are sufficiently different on each run. Ideally, one could find the boundaries of the error space, and uniformly distribute the initial conditions for the variables in the algorithm. To further process the data, the resulting matrices should be concatenated into a wide matrix,  $\mathbf{A}^M = \mathbf{A}^1 | \mathbf{A}^2 | \dots | \mathbf{A}^m$ , in which  $[\cdot]^m$  stands for the  $m$ th run of the algorithm. To remove the problem of scaling uncertainty in BSS, the concatenated mixing matrix can simply be normalized to unit variance.

The calculation of the correlation matrix  $\mathbf{R}$  should be made over all estimates. Yet, most methods are not well suited for the huge data size in fMRI. Hence, we defined the correlation coefficients based on the columns of  $\mathbf{A}^M$ . Furthermore, we are not interested in components that appear too seldomly for multiple runs of ICA. To guarantee that we focus on consistent components, a suitable threshold to  $\mathbf{R}$  is applied, and only values over it are considered.

Furthermore, if source  $\mathbf{s}_1^1$  is related to  $\mathbf{s}_1^2$ , and  $\mathbf{s}_1^2$  is to  $\mathbf{s}_4^3$ , even if  $\mathbf{s}_1^1$  is somewhat weakly related to  $\mathbf{s}_4^3$ , we may want to combine them all as estimates of a common source. As above,  $[\cdot]^m$  stands for the  $m$ th run of ICA, and  $[\cdot]_n$  for the  $n$ th component formed in that run. To do so, we raise the correlation  $\mathbf{R}$  to a power, the value of which relates to the length of such a path.

## III. METHODS

The stimulus consisted of repetitions of spoken text followed by resting periods. There were a succession of four such pairs and in each condition, 10 full head fMRI scans were

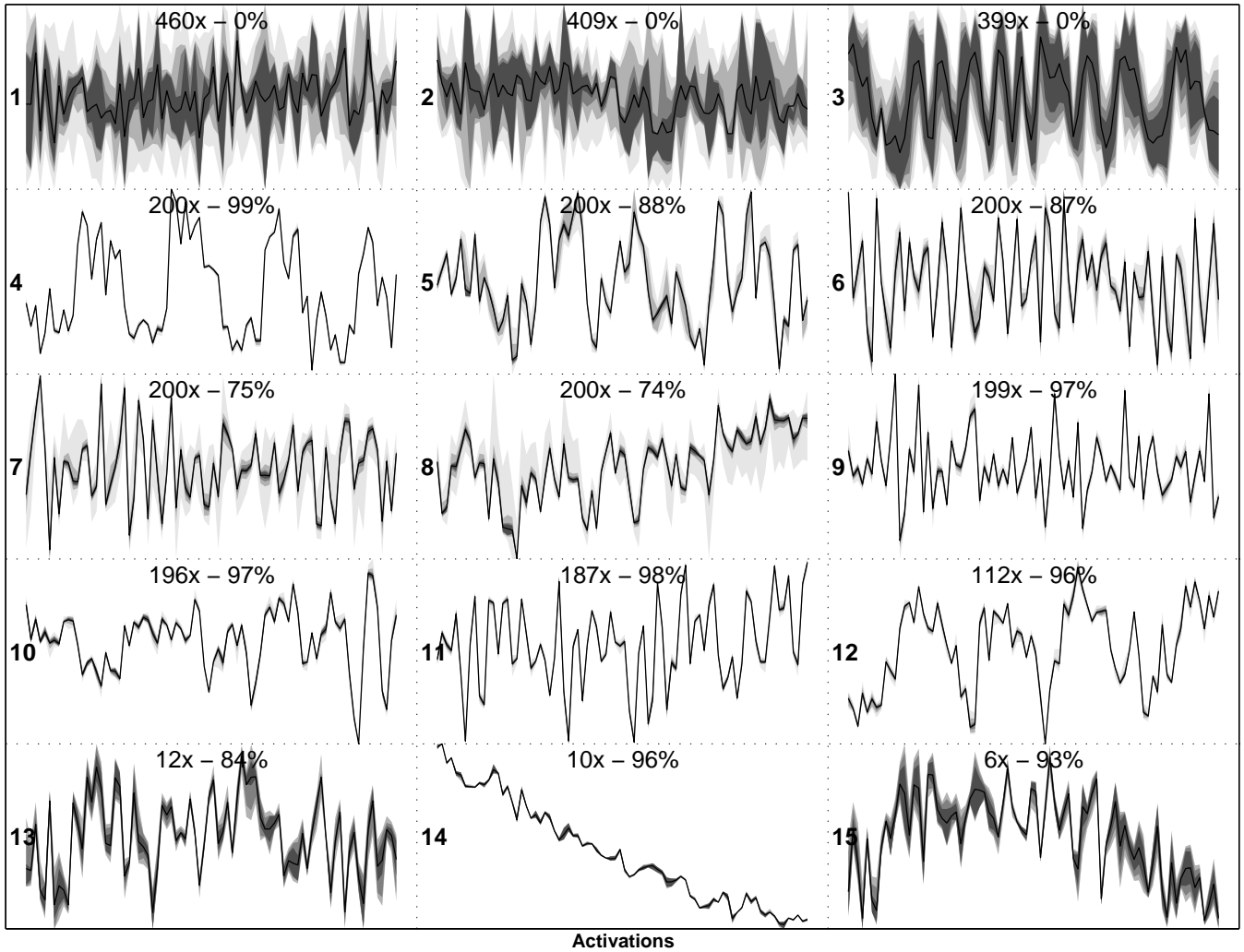


Fig. 1

OVERVIEW OF TIME COURSES OF INDEPENDENT COMPONENTS FOUND BY MULTIPLE RUNS OF ICA. THE NUMBER OF GROUPED COMPONENTS AND THE LOWEST CORRELATION ARE SHOWN ON TOP. GROUP VARIANCE IS SHOWN AS SHADES OF GRAY.

collected using a 3 Tesla GE scanner, resulting in a total of 80 volumes. The scanning repetition time was approximately 3 seconds. All volumes used T2\* settings, and consisted of  $96 \times 96 \times 37$  voxels. Data was collected at the [Advanced Magnetic Imaging Centre](#) (AMI-Centre) of the [Helsinki University of Technology](#).

After scanning, the volumes went through the usual pre-processing stages of fMRI data, which include realigning, normalizing and smoothing, as explained in [2]. These transformations compensate for head movements, allow for a comparison of volumes from different subjects and improve signal-to-noise ratio, respectively.

Because there is a considerable amount of recorded voxels external to the cortex, a suitable mask was applied to the data, which resulted in a significant decrease of non-informative data points. Furthermore, as stated earlier, the whitening stage preceding ICA included a dimension reduction, from the

original 80 volumes to the strongest 30 principal components.

FastICA was then applied 200 times, in a deflation manner, and maximizing the kurtosis contrast function. In each run, only the first 15 components were kept. From the 3000 components found in this way, the ones which reached a correlation value over 0.85 were considered similar. The thresholded values were then raised to the power 16 to account for longer paths. These numbers were chosen heuristically.

#### IV. RESULTS

Fig. 1 shows an overview of the time courses corresponding to 15 source estimates, resulting from the average of independent components grouped by correlation. Those estimates include many reliable ones, with their characteristic small variances, such as the ones depicted in frames 4, 12 and 14. Other estimates show clearly high variances, such as the ones in frames 1, 2 or 3. The number of components belonging to

each group and the smallest correlation within the group are shown on top of each frame. The black curve shows the mean time course of the group, and the gray shades correspond to the 25%, 50% and 75% quartiles of the group's distribution.

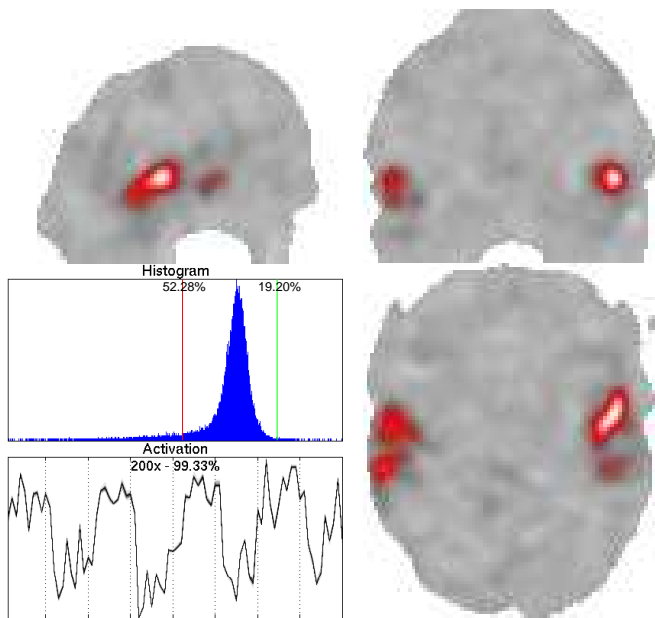


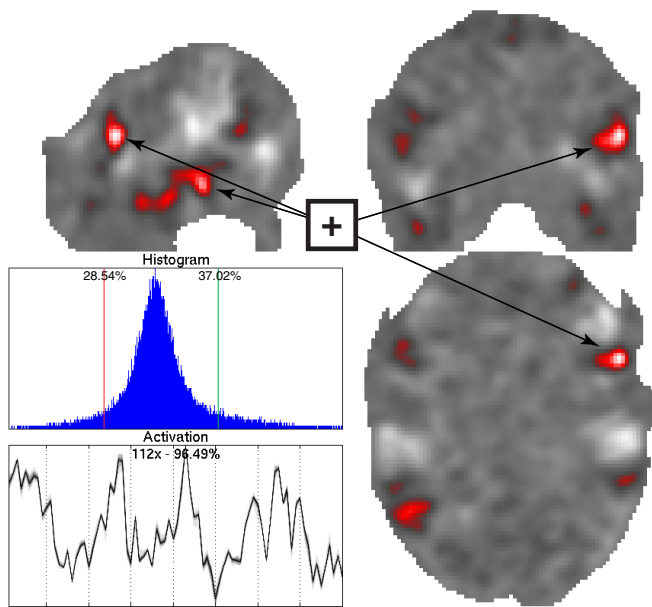
Fig. 2

SOURCE ESTIMATE REPRESENTING ACTIVATION OF PRIMARY AUDITORY AREAS.

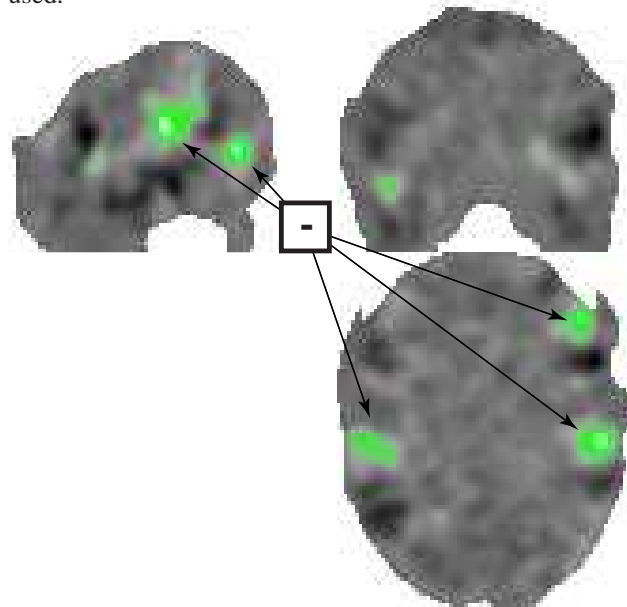
We now analyze in greater detail some of the most consistent components found in Fig. 1. The most consistent component has a time course shown in frame 4. It strongly relates with the stimulus activation pattern. This relation is further visible in Fig. 2 where vertical lines show the transitions between rest and stimulation conditions. From the three views of the estimated volume, it is clear that the activation is located laterally, close to the superior temporal gyrus, where the primary auditory cortex is located.

Furthermore, because the figures use the standard radiology orientation, it is easy to see that the left hemisphere is more active. This is all in good agreement with theory. An additional frame in Fig. 2 shows the histogram of the full estimated volume. Two vertical lines show the threshold used for visualization of the positive and negative activations. In the volume shown, no negative activation was observed.

Another consistent component, which seems to have a milder relation to the stimulation, is shown in frame 12 of Fig. 1, and in more detail in Fig. IV. Three views of the estimated volume are shown as well. In opposition to the observations in the previous estimate, here we have very significant maxima in both the positive and the negative regions of the histogram. These neighboring extremes of activation suggest contiguous areas of activation and suppression. Because of printing constraints limiting us to gray scale, we have in Fig. 3(a) only the positive activation highlighted, and in Fig. 3(b) the negative one. From the localization of the extremes, mainly around the primary auditory cortex, we can expect these to correspond to higher auditory areas, such as Broca's and Wernicke's. These are responsible for production and comprehension of speech, respectively, and fit the type of stimulus used.



(a) Positive extremes highlighted.



(b) Negative extremes highlighted.

Fig. 3 SOURCE ESTIMATE REPRESENTING ACTIVATION OF SECONDARY SPEECH RELATED AREAS.



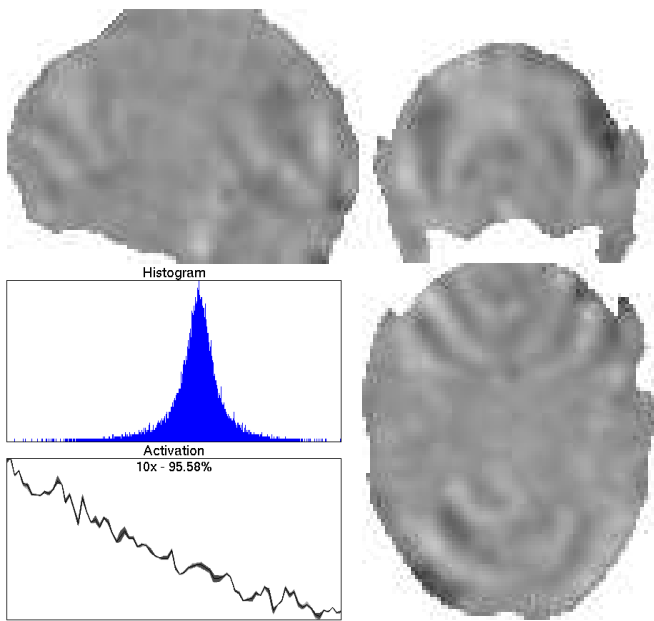


Fig. 4

SOURCE ESTIMATE WITH CLEAR ARTIFACT, PROBABLY DUE TO POOR FILTERING.

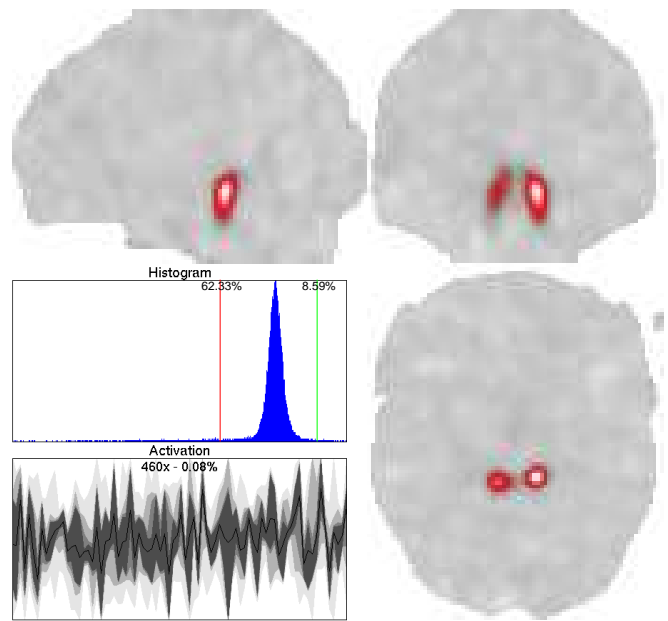


Fig. 5

SOURCE ESTIMATE POSSIBLY REPRESENTING THALAMIC ACTIVATION.

So far, one could find most of the mentioned results from a direct inspection of stimulus related components. Figure 4 gives an example of a source estimate which is hard to detect directly from the time course. In fact, the estimate evolves almost linearly in time. The estimated volume presents a clear three-dimensional ripple-like pattern, hard to accept as natural brain activation. Although the component shows great consistency, it has appeared only 10 times in the 200 runs, which suggests that it is a component hard to find even for ICA.

This clear artifact resembles the one studied in [16], using very controlled, *i.e.*, predictive settings. In practice, this type of artifacts, though commonly observed, would appear to be difficult to handle with traditional analysis methods due to a clear lack of reference.

An example of a component with high variance is shown in Fig. 5. Although the component seems unreliable, the estimated volume shows a remarkably clear activation pattern in what appears to be the thalamus. In fact, the activation pattern is somewhat posterior to the expected location for the main body of the thalamus. A possible explanation is that it may correspond to the actual location of the medial geniculate nucleus (MGN) of the thalamus. The MGN is sometimes called the auditory thalamus, due to its active involvement in relaying auditory information to the primary and secondary auditory cortices [17]. Yet, its location is often more lateral than the one observed in our study. An overlay of the activation pattern on high resolution structural MRI should improve the degree of confidence of this speculation.

The number of independent components gathered in the estimate shown in Fig. 5 is 460, which implies that more than

one independent component per run has been pulled together. How can this be, if the goal is to search close to independence? One possible explanation lies in the multiple paths that are allowed in the correlation matrix. Let us consider the simple case of a sine and a cosine. Naturally, they are uncorrelated. Yet, if we take several copies of the first, with increasing phase shifts, it is possible to build a path of high enough relations to connect both signals. Such phase jittering explanation may apply as well to frame 2 and seems very clear in frame 3 of Fig. 1.

## V. CONCLUSION

The consistency of independent component decomposition of fMRI data, in an auditory stimulation environment was tested, by exploiting the variability of the ICA algorithm used, as a function of its initialization. Using this approach, we managed to identify signal features hard to detect with methods based only on a single run.

In addition to the predicted stimulus related components, ICA has revealed activation volumes that are associated with secondary auditory areas, as well as to time varying artifacts.

We have done so in a purely spatial ICA framework. Yet, most considerations could be extended to spatio-temporal ICA algorithms, such as the ones in [18], [19].

Averaging the correlated components is just one way to estimate the correct solutions. Other methods have been used to visualize these groups of components (*c.f.* [20]), or to study the reliability of other BSS algorithms [21]. Yet, results such as the ones in the first 3 frames of Fig. 1 suggest that we may be in the presence of subspaces of solutions, rather than one-dimensional ones. The characterization of the inner structures

of the subspaces requires additional criteria other than the used decorrelation and independence. Further research should be carried out, to find better canonical representations of these spaces.

#### ACKNOWLEDGMENT

The authors would like to thank Riitta Hari and Tukka Raij from the [Advanced Magnetic Imaging Centre \(AMI-Centre\)](#) and [Brain Research Unit, Low Temperature Laboratory, Helsinki University of Technology](#) for the fMRI data and for valuable discussions on the subject of this paper.

#### REFERENCES

- [1] S. Ogawa, D. W. Tank, R. Menon, J. M. Ellermann, S. Kim, H. Merkle, and K. Ugurbil, "Intrinsic Signal Changes Accompanying Sensory Stimulation: Functional Brain Mapping with Magnetic Resonance Imaging," *Proc. Nat. Academy of Sciences*, vol. 89, pp. 51–55, 1992. [1](#)
- [2] K. J. Worsley and K. J. Friston, "Analysis of fMRI Time-Series Revisited – Again," *NeuroImage*, vol. 2, no. 3, pp. 173–235, September 1995. [1](#), [3](#)
- [3] 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, no. 3, pp. 160–188, 1998. [1](#)
- [4] M. J. McKeown, L. K. Hansen, and T. J. Sejnowski, "Independent Component Analysis of functional MRI: What Is Signal and What Is Noise?" *Current Opinion in Neurobiology*, vol. 13, no. 5, pp. 620–629, Oct 2003. [1](#)
- [5] V. D. Calhoun, T. Adali, L. K. Hansen, J. Larsen, and J. J. Pekar, "ICA of functional MRI: An Overview," in *Proc. 4th International Symposium on Independent Component Analysis and Blind Signal Separation (ICA2003)*, April 2003. [1](#)
- [6] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*. John Wiley and Sons, 2001. [1](#), [2](#)
- [7] A. Belouchrani, K. A. Meraim, J.-F. Cardoso, and E. Moulines, "Second-Order Blind Separation of Correlated Sources," in *Proc. Int. Conf. on Digital Sig. Proc.*, Cyprus, 1993, pp. 346–351. [1](#)
- [8] A. Ziehe and K.-R. Müller, "TDSEP – An Effective Algorithm for Blind Separation Using Time Structure," in *Proc. Int. Conf. on Artificial Neural Networks (ICANN'98)*, Skövde, Sweden, 1998, pp. 675–680. [1](#)
- [9] A. Hyvärinen and E. Oja, "A Fast Fixed-Point Algorithm for Independent Component Analysis," *Neural Computation*, vol. 9, pp. 1483–1492, 1997. [2](#)
- [10] "The FastICA MATLAB Package," Available for download at <http://www.cis.hut.fi/projects/ica/fastica/>. [2](#)
- [11] J. Särelä and R. Vigário, "Overlearning in Marginal Distribution-Based ICA: Analysis and Solutions," *Journal of Machine Learning Research*, vol. 4, pp. 1447–1469, Dec 2003. [2](#)
- [12] M. J. McKeown and T. J. Sejnowski, "Independent Component Analysis of fMRI Data: Examining the Assumptions," *Human Brain Mapping*, vol. 6, pp. 368–372, 1998. [2](#)
- [13] T. P. Jung, S. Makeig, M. J. McKeown, A. J. Bell, T. W. Lee, and T. J. Sejnowski, "Imaging Brain Dynamics Using Independent Component Analysis," in *Proceedings of the IEEE*, vol. 89, no. 7, July 2001, pp. 1107–1122. [2](#)
- [14] R. Vigário, J. Särelä, V. Jousmäki, M. Hämäläinen, and E. Oja, "Independent Component Approach to the Analysis of EEG and MEG Recordings," *IEEE Trans. Biomed. Eng.*, vol. 47, no. 5, pp. 589–593, May 2000. [2](#)
- [15] F. Esposito, E. Formisano, E. Seifritz, R. Goebel, R. Morrone, G. Tedeschi, and F. D. Salle, "Spatial Independent Component Analysis of functional MRI Time-Series: To What Extent Do Results Depend on the Algorithm Used?" *Human Brain Mapping*, vol. 16, no. 3, pp. 146–157, Jul 2002. [2](#)
- [16] A. M. Smith, B. K. Lewis, U. E. Ruttimann, F. Q. Ye, T. M. Sinnwell, Y. Yang, J. H. Duyn, and J. A. Frank, "Investigation of Low Frequency Drift in fMRI Signal," *NeuroImage*, vol. 9, no. 5, pp. 526–533, May 1999. [5](#)
- [17] J. Rademacher, U. Bürgel, and K. Zilles, "Stereotaxic Localization, Intersubject Variability, and Interhemispheric Differences of the Human Auditory Thalamocortical Systems," *NeuroImage*, vol. 17, pp. 142–160, 2002. [5](#)
- [18] K. Suzuki, T. Kiryu, and T. Nakada, "Fast and Precise Independent Component Analysis for High Field fMRI Time Series Tailored Using Prior Information on Spatiotemporal Structure," *Human Brain Mapping*, vol. 15, no. 1, pp. 54–66, Jan 2002. [5](#)
- [19] J. V. Stone, J. Porrill, N. R. Porter, and I. D. Wilkinson, "Spatiotemporal Independent Component Analysis of Event-Related fMRI Data Using Skewed Probability Density Functions," *NeuroImage*, vol. 15, no. 2, pp. 407–421, Feb 2002. [5](#)
- [20] J. Himberg and A. Hyvärinen, "Icasso: Software for Investigating the Reliability of ICA Estimates by Clustering and Visualization," in *In Proc. 2003 IEEE Workshop on Neural Networks for Signal Processing (NNSP2003)*, Toulouse, France, 2003, pp. 259–268. [5](#)
- [21] F. Meinecke, A. Ziehe, M. Kawanabe, and K.-R. Müller, "A Resampling Approach to Estimate the Stability of One-Dimensional or Multidimensional Independent Components," *IEEE Trans. Biomed. Eng.*, vol. 49, no. 12, pp. 1514–1525, 2002. [5](#)