

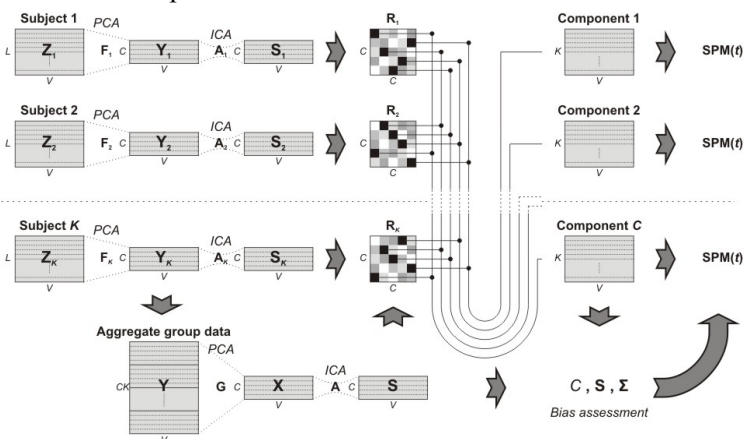
Unbiased group-level statistical assessment of independent component maps by means of automated retrospective matching

D. Langers^{1,2}

¹Otorhinolaryngology, University Medical Center Groningen, Groningen, Netherlands, ²Eaton-Peabody Laboratory, Massachusetts Eye and Ear Infirmary, Boston, MA, United States

Introduction. – In the context of fMRI data analyses, spatial Independent Component Analysis (sICA) has become commonplace in recent years [1]. Its primary potential advantage over conventional mass-univariate regression methods is that it does not require prior specification of hemodynamic response functions, and can thus be applied in settings where response dynamics are unknown (e.g. in experiments with complex tasks/stimuli, or in resting state designs). However, because sICA is primarily an exploratory technique that requires retrospective interpretation, statistical assessments are difficult, especially at the group level. In literature, temporal concatenation followed by back-projection appears to be the most popular approach [2], but so far its behavior under the null hypothesis and its potential susceptibility to bias has not been addressed. In this study, an alternative group-level sICA method is presented and validated.

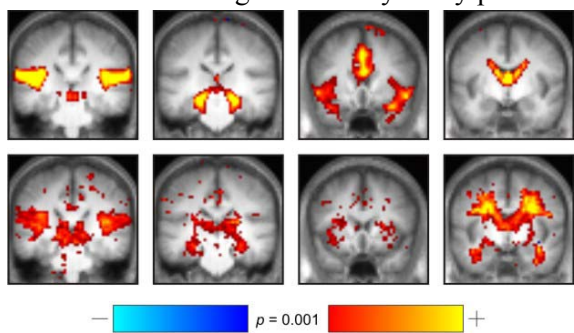
Theory. – The proposed method is outlined in the right figure. Initially, individual subject-level analyses are performed, each consisting of a principal components data reduction step (PCA) followed by sICA. Next, the reduced subject data are concatenated to a single group-level aggregate dataset, which is reduced with PCA and decomposed with sICA as well. Then, the inner product matrices between each of the subject- and group-level component maps are used to identify optimal individual one-on-one pairings by means of the Hungarian (Kuhn-Munkres) algorithm [3,4]. Finally, all subject component maps that are matched to the same group component map are entered into a group-level statistical analysis.



Although the aforementioned pairing process may introduce a non-zero selection bias, the magnitude of this bias can be predicted analytically and corrected for. A detailed description is beyond the scope of this abstract [5].

Materials & Methods. – Various simulations were run to test the validity and performance of the proposed method. Null data consisting of Gaussian noise images without any effects of interest were generated to validate the method under the null hypothesis. In addition, artificial data volumes with true effects were generated to evaluate the method’s sensitivity, specificity, and discriminatory power. Finally, an available fMRI dataset of 32 subjects that involved a basic audiovisual task were analyzed to assess the practicality of the method.

Results. – The validity of the method was assessed using artificial null data. Bias effects proved significant when not corrected for: false positive rates up to 0.34 were observed when thresholding at $P = 0.05$. However, when implementing a correction according to the analytically predicted bias, no deviations from the imposed confidence levels were found. The results from artificial datasets with true effects showed that the proposed method was marginally conservative, and further corroborated that confidence levels were never exceeded. Moreover, it showed that the method reached acceptable discriminatory power (e.g.: $d' = 3.52$ where an uncorrected method reached $d' = 3.68$). Finally, actual fMRI outcomes showed that various meaningful components in the aggregate dataset could also be detected consistently in individual subjects.



The left figure shows normalized aggregate template maps (top) and bias-corrected group-level statistical parametric maps (bottom) for four components (Auditory, Limbic, Insula+ACC, and Ventricular CSF).

Conclusions. – The proposed bias-corrected method was shown to result in valid outcomes, while its power was only moderately worse than a biased method. In practice, the method succeeded in extracting significant effects of interest. Yet, bias effects were shown to be considerable when not corrected for; this underlines the necessity for a proper statistical characterization of group-level ICA methods under the null hypothesis, which -so far- is lacking for other methods.

References. – [1] Calhoun et al. *NeuroImage* (2009) 45(1): S163-72; [2] Calhoun et al. *HBM*. (2001) 14(3): 140-51; [3] Esposito et al. *Magn.Res.Imag.* (2008) 26(7): 905-13; [4] Kuhn. *Nav.Res.Log.Quart.* (1955) 2: 83-97; [5] Langers. *HBM*. (In press).