Lab Report

Hanna Röhling

April 4, 2016

Contents

1 Introduction

The analysis of functional magnetic resonance imaging (fMRI) data by means of multivariate pattern analysis gives rise to the possibility to at least partially decode and

reconstruct mental states. When exploring computational methods for decoding, a focus on the human visual cortex seems natural, since its structure and properties are relatively well studied and visual stimuli are easily produced and controlled.
In this project Oliver Eberle and I investigated multivariate decoding methods on simulated and real fMRI data of the human visual cortex. The real fMRI data used were acquired in an experiment by Heinzle et al. (2011) that showed "that fMRI signal fluctuations reveal a detailed retinotopically organized functional connectivity structure between the visual field maps of remote areas of the human visual cortex" [2]. The generated visual stimulus and corresponding voxeldata used in the simulation experiments were also based on the study of Heinzle et al. with respect to their properties (i.e. duration, stimulus intensities, etc.). Therefore one must keep in mind that at least the real dataset was originally not intended to be used in the context of decoding and reconstruction of visual stimuli.
At first we examined the possibilities and limits of the decoding toolbox [1]. In order to do so we simulated data on the basis of a flickering checkerboard disc stimulus. The simulated dataset was generated by applying different mapping functions between stimulus and voxeldata as well as varying signal to noise ratios. After those first steps Oliver focused on the decoding of finger presses and the estimation of a retinotopic model (cf. Oliver's lab report). I on the other hand focussed on the estimation of an accurate chancelevel for the experiment using simulated data. Furthermore I estimated the ideal delay between voxeldata-stimulus pairs that went into the decoding analysis using the real dataset of Heinzle et al. [2].

2 Data and the Decoding Toolbox

As mentioned above, the data used in our project was originally part of a functional connectivity study with respect to the human visual cortex. Heinzle et al. describe the experimental setup (S+) as follows:

"[...] subjects were engaged in a fixation task (Landolt C) at the center of the visual display. Visual stimuli were projected onto a translucent screen at the rear of the scanner (screen size: 25 by 20 of visualangle). During S+, subjects were presented with 8 runs of a highcontrast, flickering circular checkerboard subdivided into 48 segments which changed their local contrast (4 logarithmic levels between 0.1 and 1) independently over time [...]. Note that the visual stimulus had a border along the vertical meridian resulting in completely independent visual stimulation in the two visual hemifields. [...]" [2].

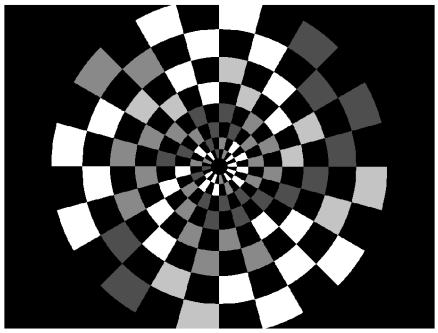

Figure 1: Example frame of checkerboard stimulus

See Heinzle et al. [2] for further information on participants, fMRI image acquisition and data preprocessing.
For the multivariate analyses of the real and simulated fMRI data we mostly employed "The Decoding Toolbox" (TDT) [1], which allows for different voxel selection criteria such as searchlight, region-of-interest (ROI) and wholebrain analyses. !!!!!!!!!!!!!!!!!!!!!!

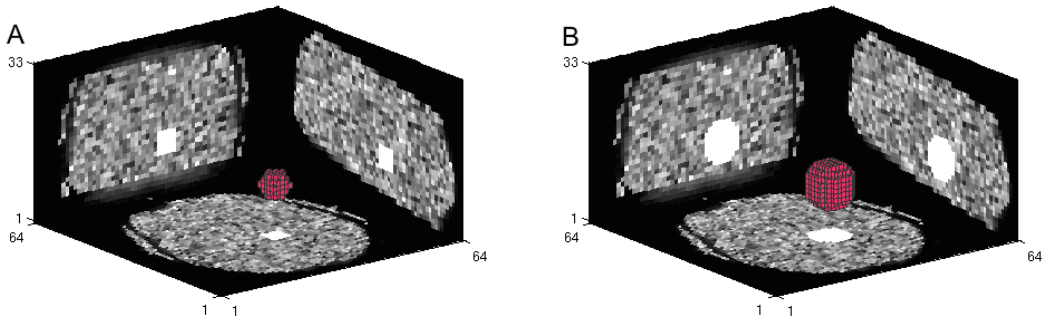

Figure 2: Example searchlights of size 3 (A) and size 5 (B).

3 Toydata Decoding and Reconstruction

First we wanted to determine whether the TDT was capable of reconstructing a visual stimulus from toydata that was generated in a very simple manner.

$$V = MS, \tag{1}$$

with $M \in \mathbb{R}^{d \times n}$ being the mapping matrix, $V \in \mathbb{R}^{d \times t}$ being the voxeldata over time and $S \in \mathbb{R}^{n \times t}$ the stimulus over time.
First we defined M to perform a direct mapping between stimulus and voxeldata, meaning that $\forall\ v_{i,*}\ i \in \{1, ..., d\}\ \exists\ j \in \{1, ..., n\}$: $v_{i,*} = s_{j,*}$. Hence, $\forall\ i \in \{1, ..., d\}$, $j \in \{1, ..., n\}$: $m_{i,j} \in \{0, 1\}$ and $\forall i \in \{1, ..., d\}$: $\sum_{j=1}^{n} m_{i,j} = 1$. !!!!!!!!!!!!!!!!!!!!!!

4 Chancelevel Estimation

To estimate the chance-level of the decoding accuracy, I generated a toy-stimulus corresponding to the task used in the experiment by Heinzle et al.. It also consists of 8 different runs à 100 time points. In this case the stimulus has only one sector that changes its intensity over time (i.e. it can attain the value 1, 2, 3 or 4). Consequently we assumed the chancelevel to be 25%. The intensities are balanced with respect to each run, meaning that the stimulus takes on each of the four values exactly 25 times within each run. For the chance-level estimation I generated voxeldata of different dimensionalities (1, 2, 10, 100, 500 and 1000) that randomly took on values from the standard

normal distribution. I then performed a decoding analysis on the generated voxeldata by means of classification and regression. This simulation procedure was repeated 10000 times for each of the six dimensionalities. For both decoding methods I looked at the distribution of Pearson correlation, Fisher-z-transformed correlation, "overall Pearson correlation" and "overall Spearman correlation". The difference between the correlation computed within TDT [**?**] and the "overall correlation" (Pearson and Spearman) mentioned at last lies in when and with what data the respective values are acquired. In the toolbox the correlation is computed by taking the mean value of the correlation between true and predicted labels of the test data chunk within each step of the cross validation. Hence it is the mean of 8 different values that correspond to the correlation of true and predicted labels of 100 points in time. The "overall correlation" however, is the correlation between the final predicted labels and the true labels. Therefore the "overall correlation" corresponds to the correlation of 800 points in time. The accuracy distribution was taken into account for the classification only. For the regression approach it would be virtually zero, because the predicted label would hardly ever exactly match the discrete true label. For reasons of clarity and comprehensibility I only included the voxel dimensions 1, 2 and 100 in Figures 3 and 4.

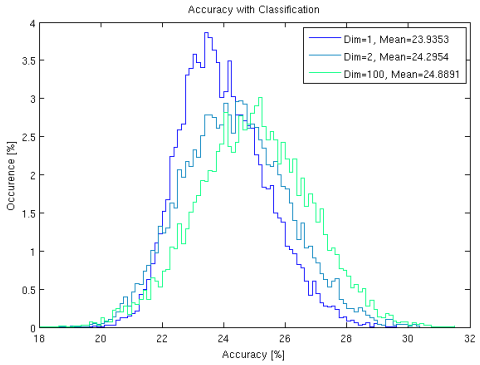

Figure 3: Histogram of decoding accuracies for the 10000 simulations of the chancelevel estimation.

At least by visual inspection the correlation distributions resulting from the chance level estimation seem to differ significantly from the distribution of Pearson's correlation coefficient (approximated by the t-distribution). In order to examine how exact Pearson's correlation is in this situation I compared the empirical p-values and the p-values of the distribution of Pearson's correlation (figures 5 to 6). The empirical p-values can be computed directly from the distributions shown in figure 4.

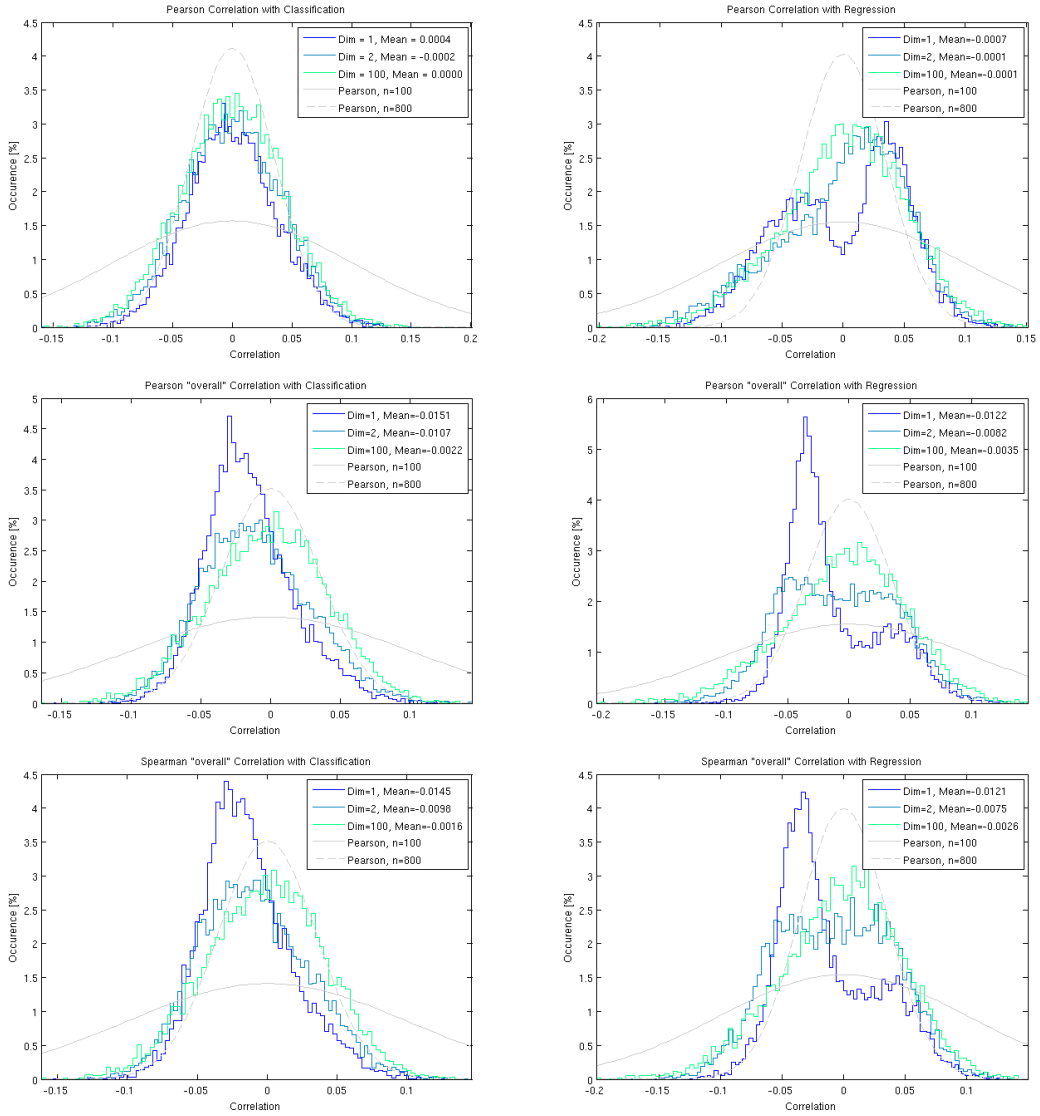

Figure 4: Correlation distributions for the 10000 simulation runs of the chancelevel estimation.

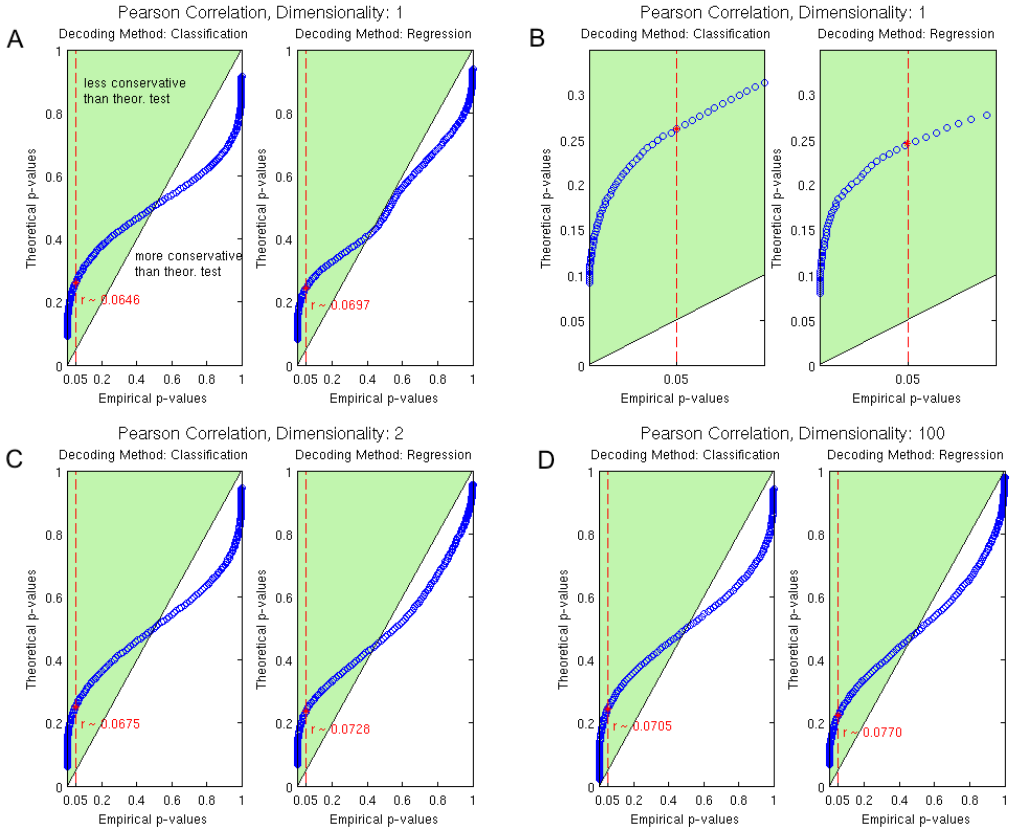

Figure 5: Direct comparison of empirical and theoretical p-values for the Pearson correlation for classification and regression, dimensionality of voxeldata: 1 (A), 2 (C) and 100 (D). In B a close up of the graph in A is depicted. For a closer look at dimensionalities 2 and 100 see figure **??** in the appendix.

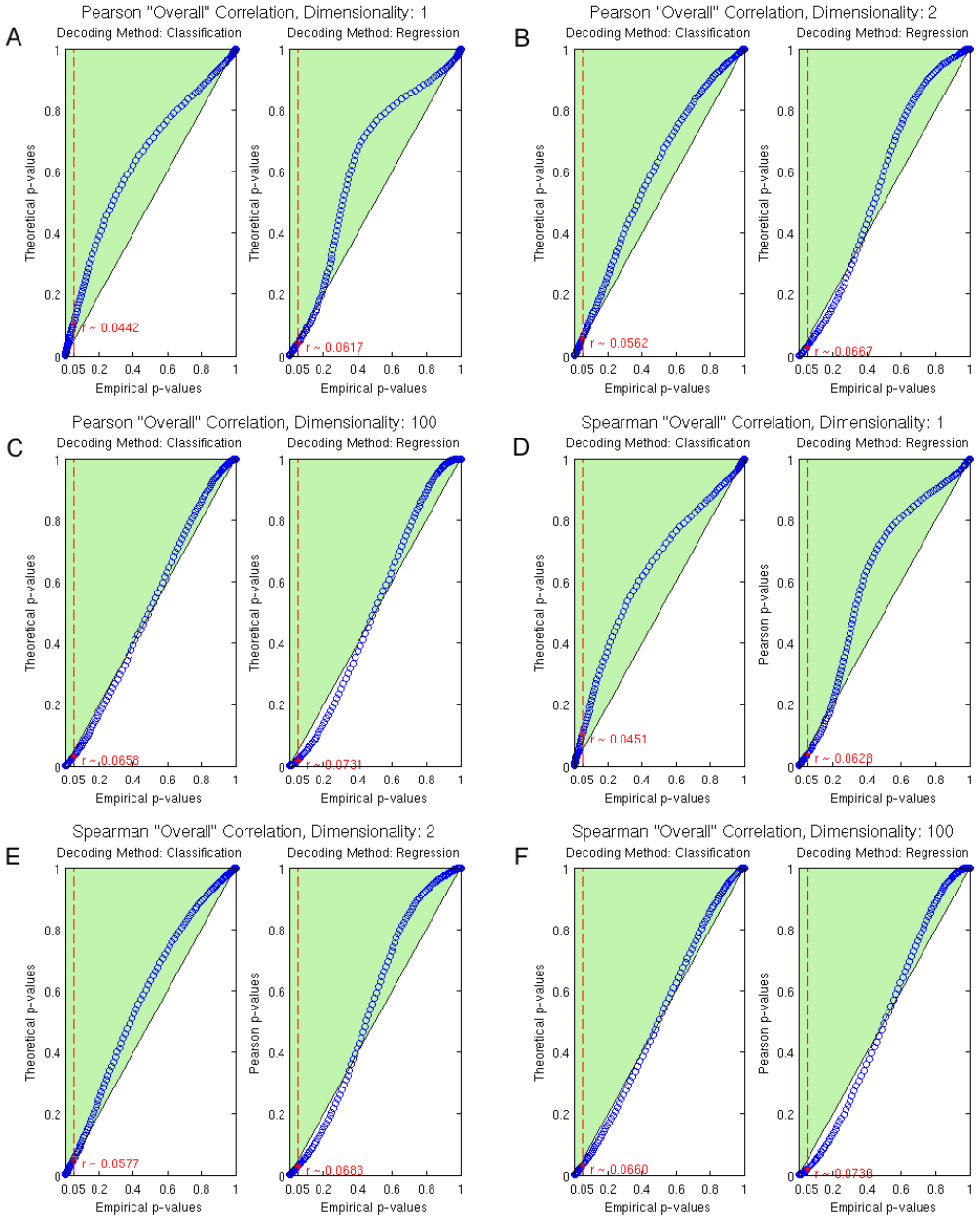

Figure 6: Direct comparison of empirical and theoretical p-values for the "overall" Pearson (A, B, C) and Spearman (D, E, F) correlation for classification and regression, dimensionality of voxeldata: 1, 2 and 100. For a closer look around the empirically determined critical value of 0.05 see figure 16.

5 Choice of Delay

The BOLD response is a relatively slow measurement of brain activity, since the underlying mechanism is hemodynamics. Because of that one has to introduce a delay in order to get "correct" stimulus and response pairings. However, when looking at a typical hemodynamic response function (HRF) curve it is not obvious which delay to choose, since the response to a stimulus takes several seconds. Furthermore the HRF can vary across different brain regions and subjects and is dependent on the presented stimulus. In order to find an appropriate delay I ran several ROI analyses with delays 0 ms, 1500 ms, 3000 ms, 4500 ms, 6000 ms and 7500 ms. All of these delays are multiples of TR and therefore result in the selection of different images from the fMRI data for different delays in each case. In order to compare them I took the mean accuracy over subjects and twelve ROIs (V1, V2A, V2B, V3A, V3B, V4 each in both hemispheres) for each of the 48 patches of the checkerboard stimulus. In a pairwise comparison of delays I used paired t-tests to determine whether the means of the distribution differ significantly. A delay of 3000 ms gave significantly better accuracies averaged over subjects and ROIs than all other delays, as can be seen in figure 7. Pairwise comparisons between all other delays can be found in the supplemental material section (figure 19).
Despite the fact that a delay of 3000 ms yields significantly better results than the others, the patches decoding accuracies still seem to be pretty close to the chancelevel of about 25 %. With the averaging procedure described above a lot of information gets lost that actually underlines that the decoding is successful and not mere "guessing" of the classifier. When taking a closer look at the accuracies for only V1 and for both hemisphere's individually one can see that the accuracy for the superior delay 3000 reach accuracies of up to 50 %. Furthermore they change periodically. This can be explained by the fact that visual stimulus information from the left hemifield is at first projected to the right hemisphere and vice versa. By projecting the decoding accuracies onto the checkerboard stimulus (figure 9) this become more clear. If decoding is performed with the ROI V1 in the left hemisphere the patches in the right visual field have a higher accuracy and the other way around. In figures 8 and 9 this is depicted for the ROI V1 in the left hemisphere. The respective figures for the left hemisphere can be found in figures 20 and 21 in the appendix.

6 Discussion

!!!!!!!!!!!!!!!!!!!!!!

References

[1] M. N. Hebart, K. Görgen, and J.-D. Haynes. The decoding toolbox (tdt): a versatile software package for multivariate analyses of functional imaging data. *Frontiers in Neuroinformatics*, 8(88), 2015.

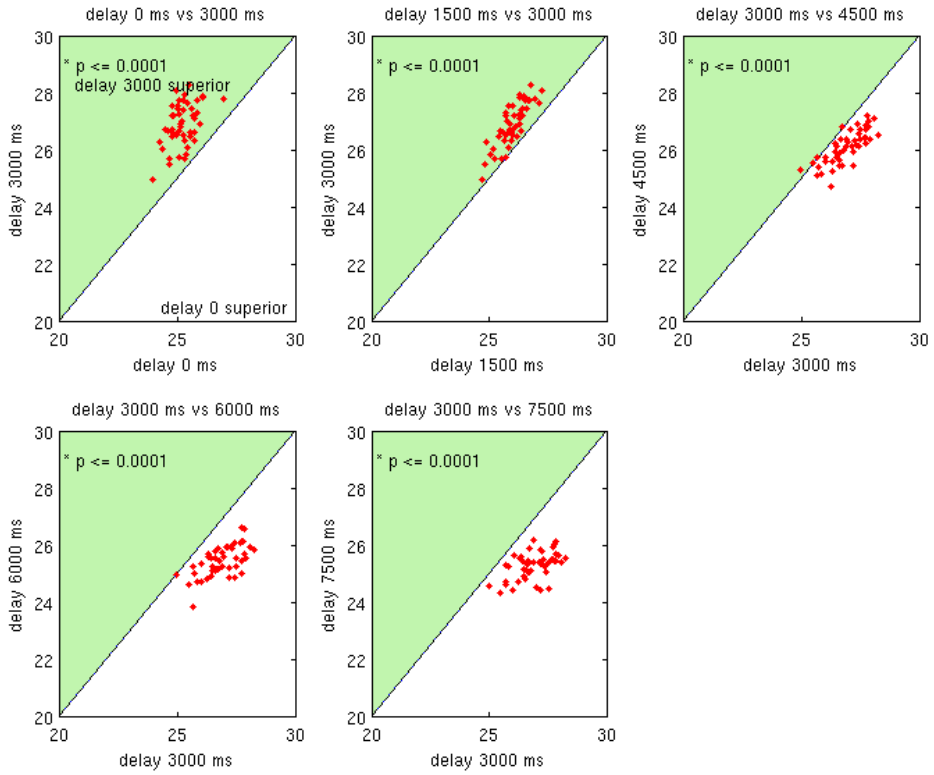

Figure 7: Pairwise comparison of accuracies for different delays for each of the 48 patches of the checkerboard stimulus (averaged over four subjects and 12 ROIs). A star (*) means that the distributions differ significantly (t-test).

[2] J. Heinzle, T. Kahnt, and J.-D. Haynes. Topographically specific functional connectivity between visual field maps in the human brain. *NeuroImage*, 56(3):1426–1436, 2011.

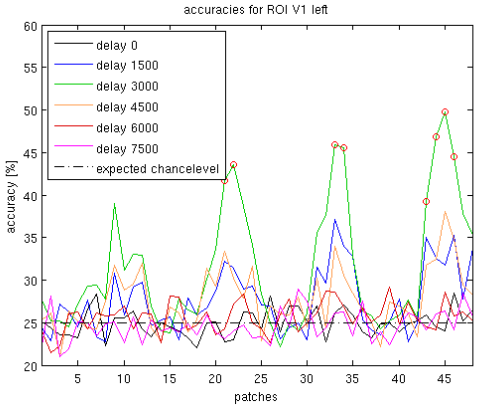

Figure 8: Decoding accuracies for different delays averaged over subjects for ROI V1 in the left hemisphere, circled in red are the results for the eigth patches that yield the best results.

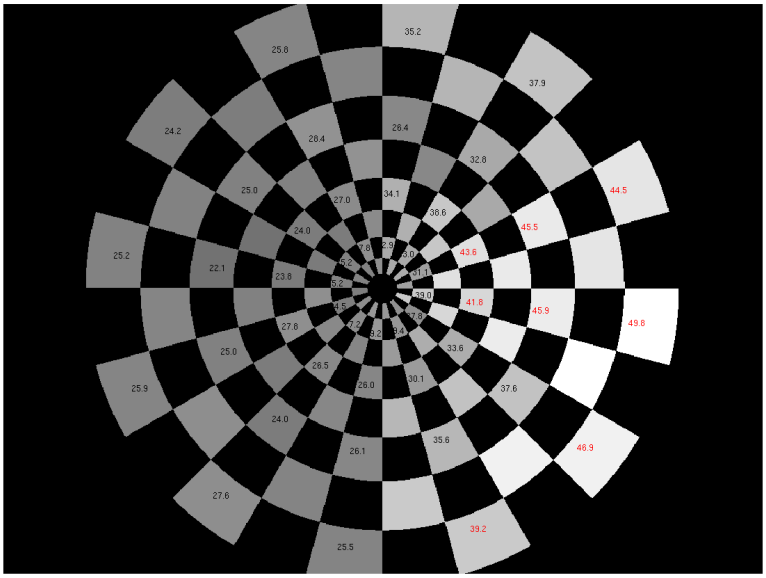

Figure 9: Averaged (over subjects) decoding accuracies for delay 3000 with ROI V1 of the left hemisphere, written in red are the results for the eight patches that yield the best results.

A Appendix

A.1 Important Scripts and Functions

A.1.1 Toydata decoding and reconstruction

- decode_simulated_data_new.m
- decode_simulated_data_new_regression.m
- ultimate_toydata !!!! BERARBEITEN
- toyvoxel_mapping_diff_dim.m

A.1.2 Chancelevel Estimation

-

A.1.3 Choice of Delay

- home/scripts_for_Kai/compare_delays_pw.m
- home/scripts_for_Kai/compare_Delays.m
- plot_checkerboard_mark_best.m (modified from Oliver's plot_checkerboard.m)

A.1.4 SVM Weights versus Patterns

- transres_SVM_pattern.m (modifications)
- transres_SVM_weights.m (modifications)
- plot_SVMweights_vs_patterns.m

A.2 Supplemental Figures

A.2.1 Chancelevel Estimation

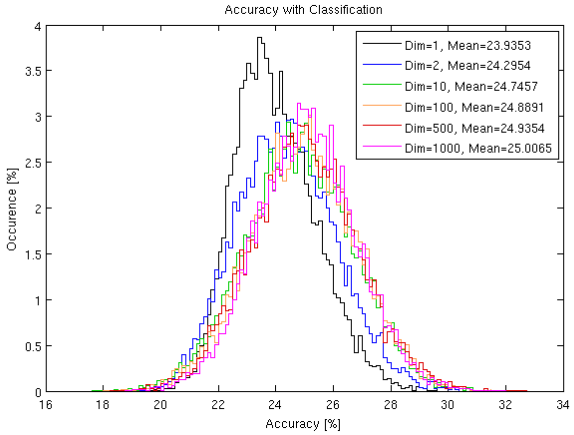

Figure 10: Histograms of the decoding accuracies in 10000 simulations with respect to classification decoding for different toyvoxel dimensionalities (1,2,10,100,500,1000).

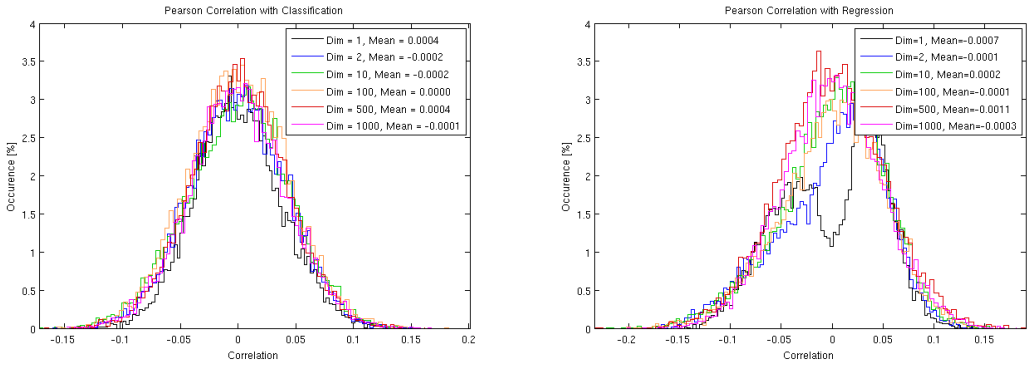

Figure 11: Histograms of the correlation between .....

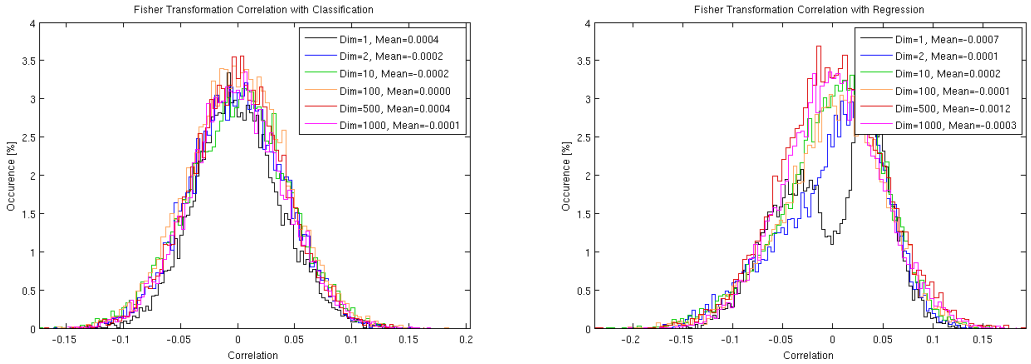

Figure 12: CAPTION.

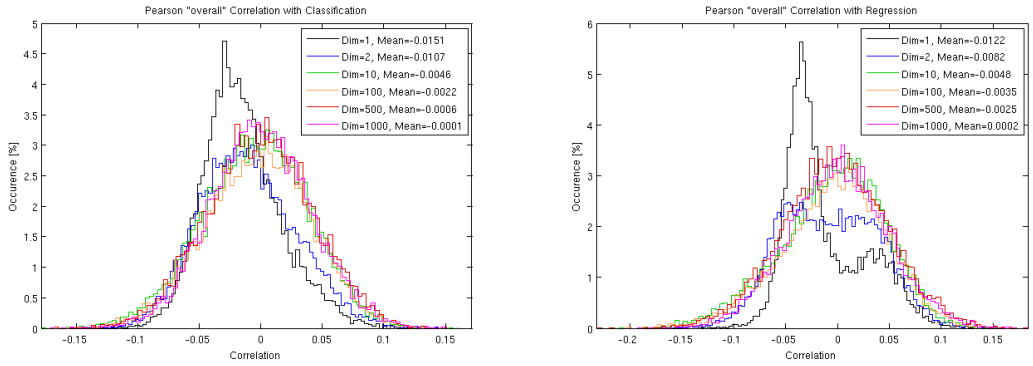

Figure 13: CAPTION.

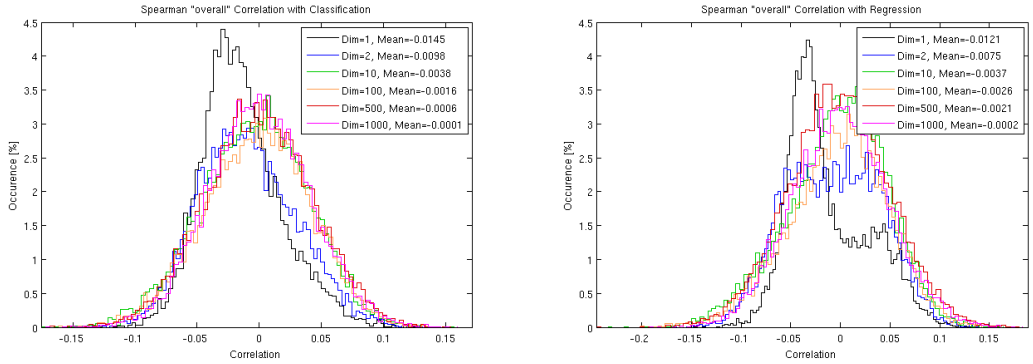

Figure 14: CAPTION.

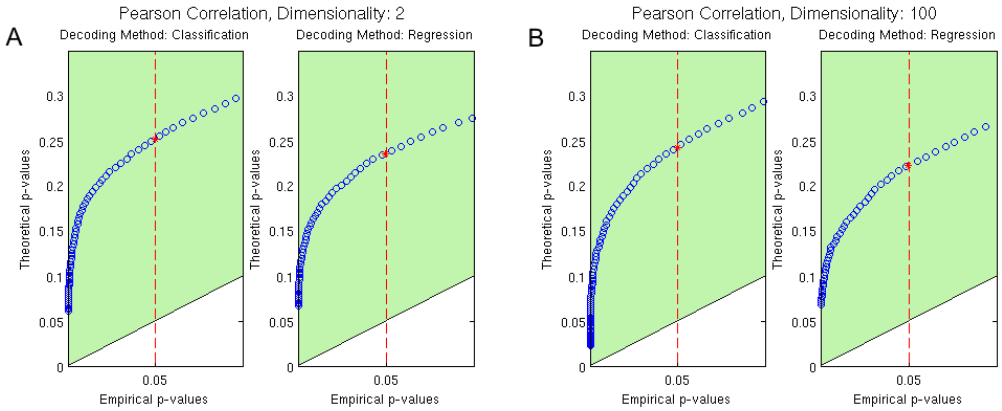

Figure 15: A close-up of the direct comparison of empirical and theoretical p-values around the empirical value for 0.05 for the Pearson correlation for classification and regression, dimensionality of voxeldata: 2 (A) and 100 (B).

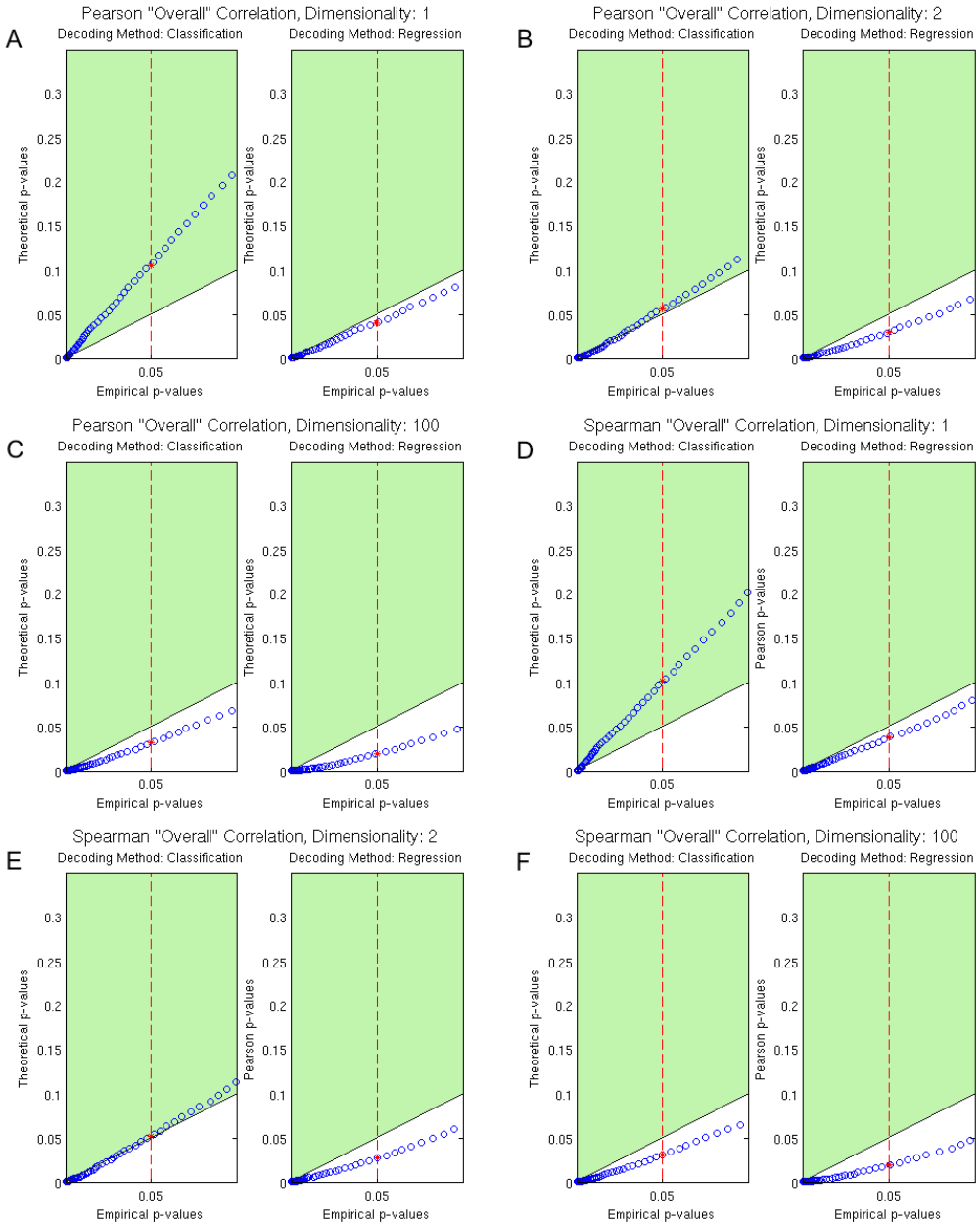

Figure 16: A close-up of the direct comparison of empirical and theoretical p-values around the empirical value for 0.05 for the "overall" Pearson (A, B, C) and Spearman (D, E, F) correlation for classification and regression, dimensionality of voxeldata: 1, 2 and 100.

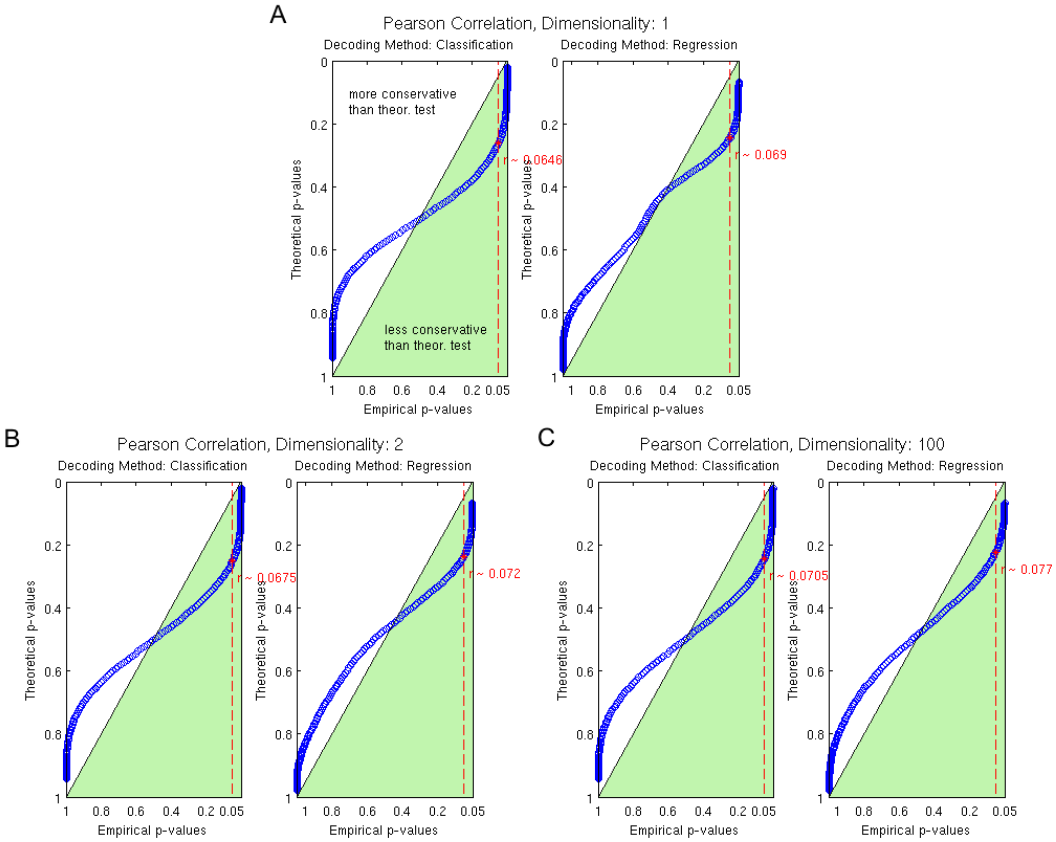

Figure 17: OTHERSIDE.

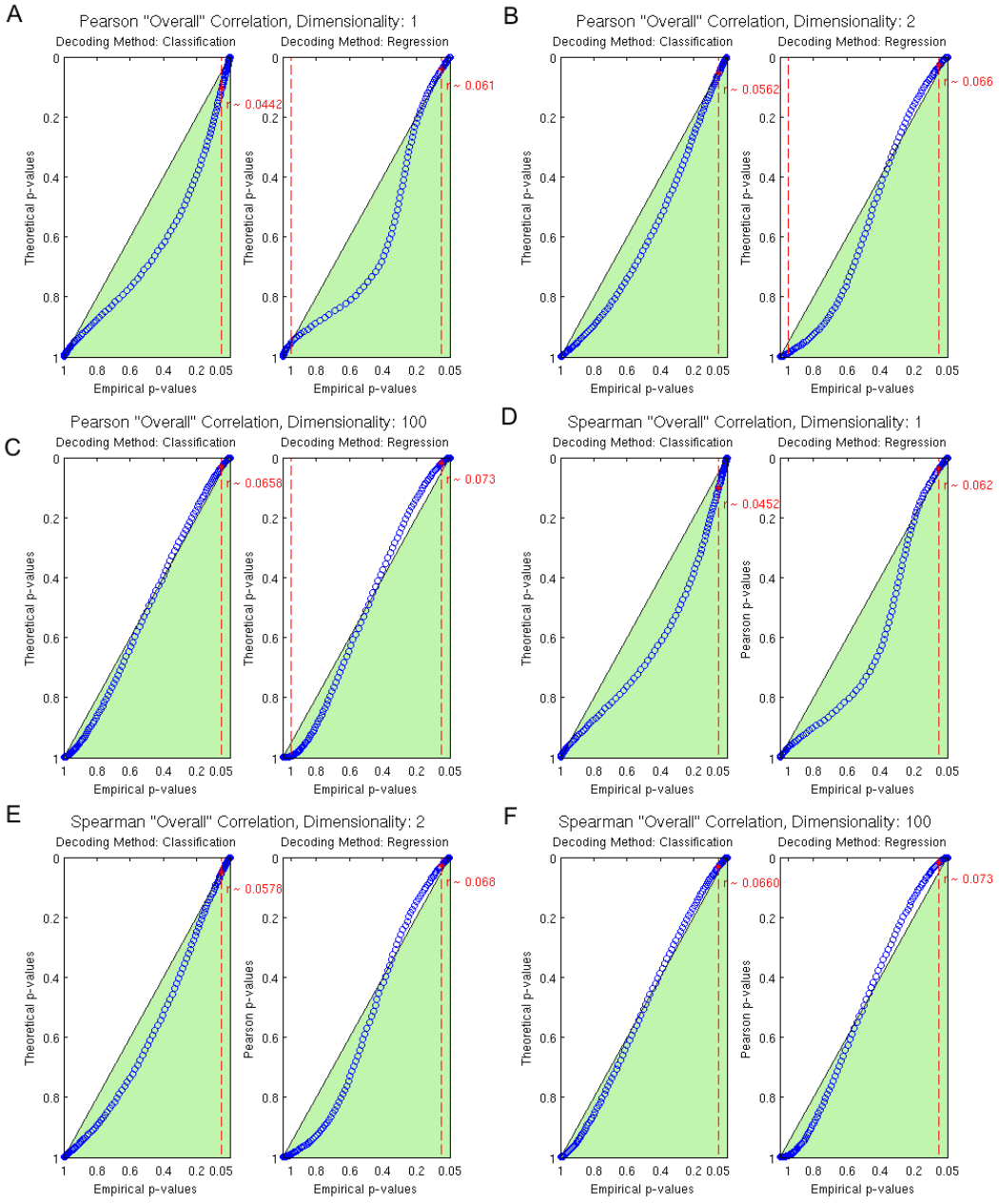

Figure 18: OTHERSIDE TEST.

A.2.2 Choice of Delay

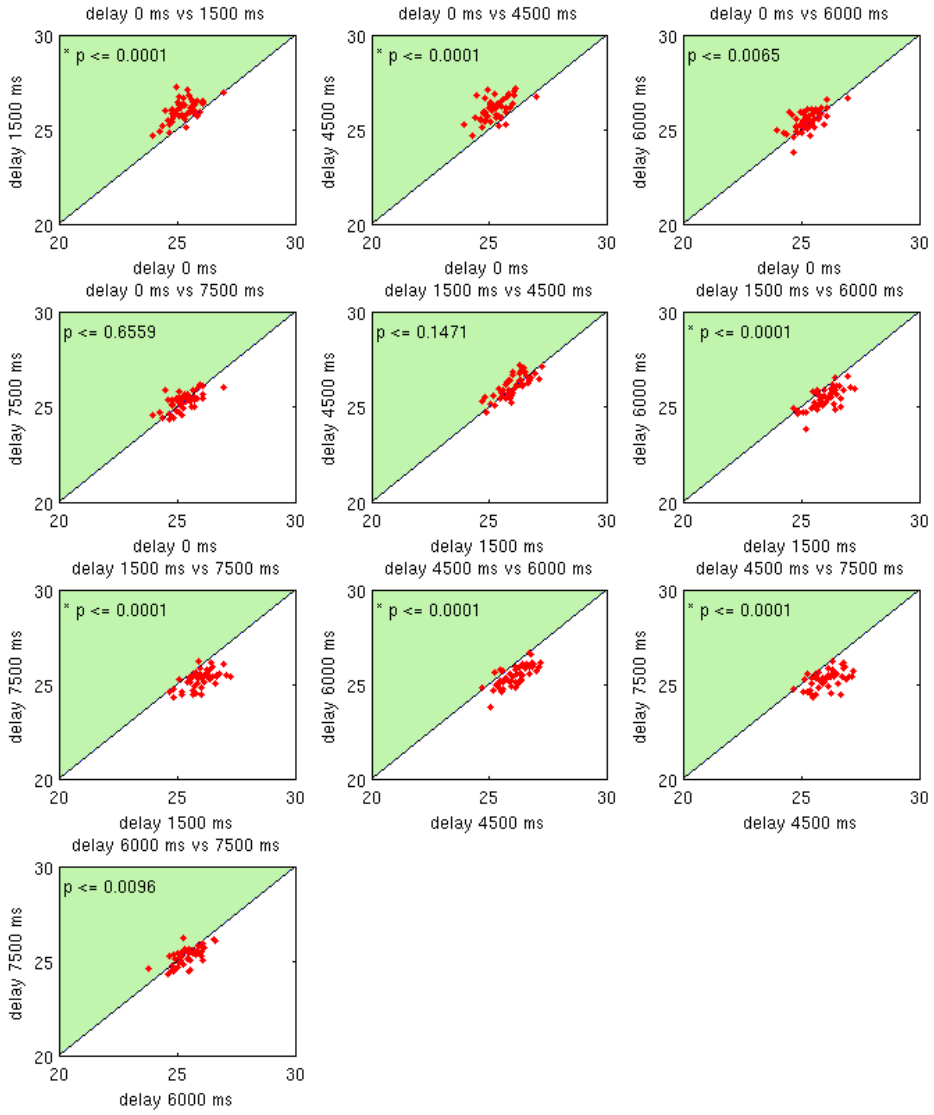

Figure 19: Pairwise comparison of delays for each patch of the checkerboard stimulus (averaged over subjects and ROIs). A star (*) means that the distributions differ significantly (t-test).

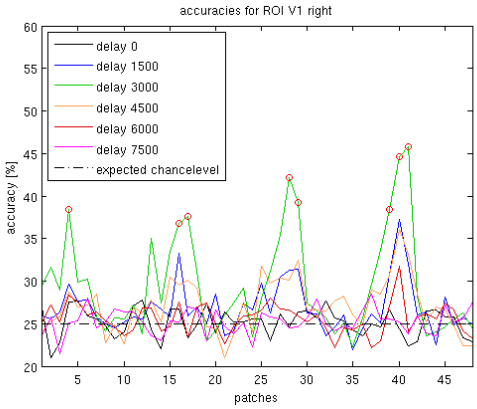

Figure 20: Decoding accuracies for different delays averaged over subjects for ROI V1 in the right hemisphere, circled in red are the results for the eigth patches that yield the best results.

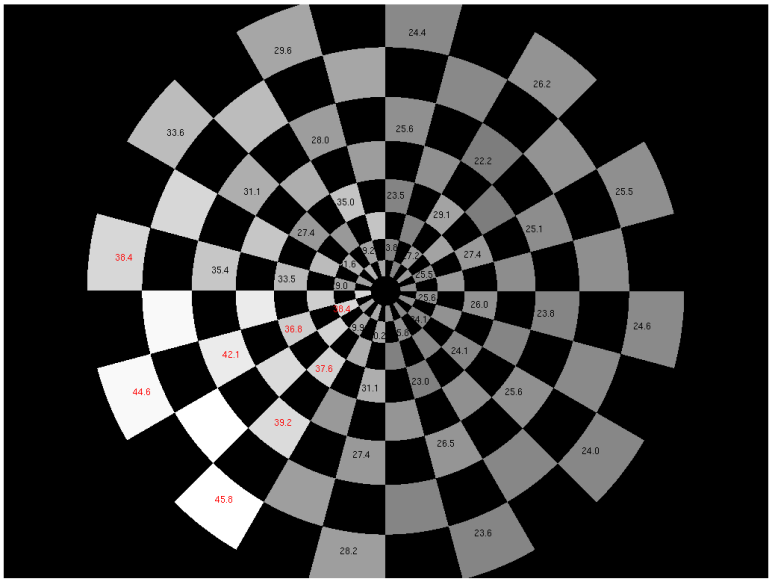

Figure 21: Averaged (over subjects) decoding accuracies for delay 3000 with ROI V1 of the right hemisphere, written in red are the results for the eight patches that yield the best results.

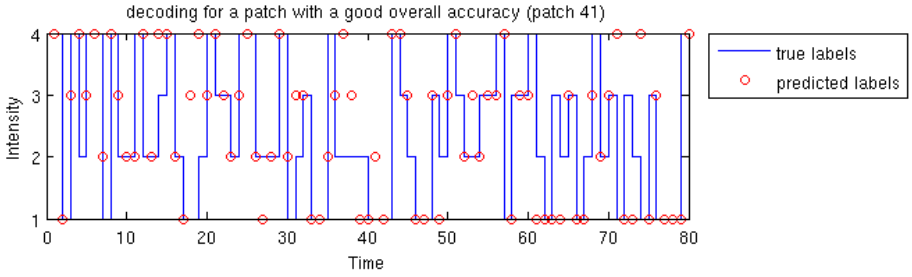

Figure 22: Predicted and true labels for decoding with the ROI V1 of the right hemisphere for patch 41 which yielded relatively good results in the decoding accuracies over subjects (see figure 20).

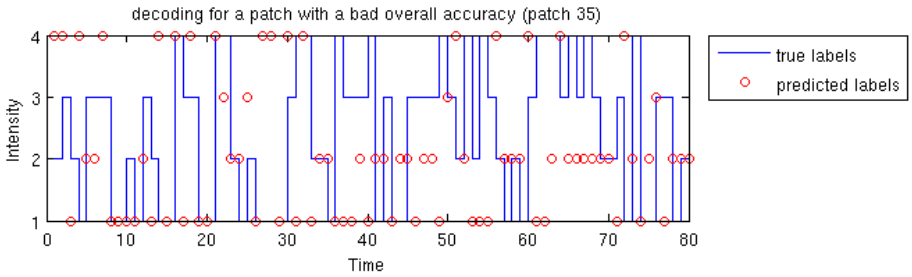

Figure 23: Predicted and true labels for decoding with the ROI V1 of the right hemisphere for patch 35 which yielded relatively bad results in the decoding accuracies over subjects (see figure 20).

A.2.3 SVM Weights versus Patterns

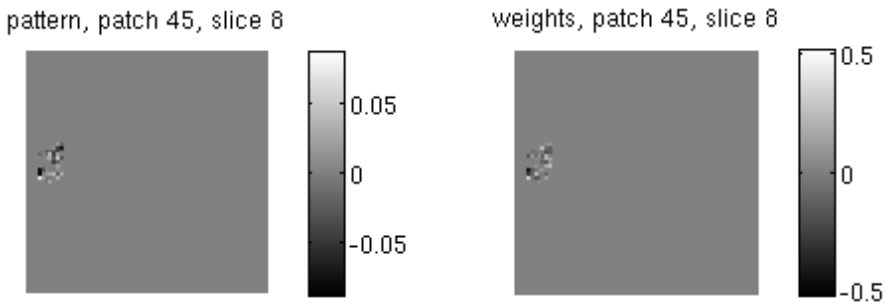

Figure 24: Comparison of patterns and weights of subject 7 in V1 (slice 8) [2].

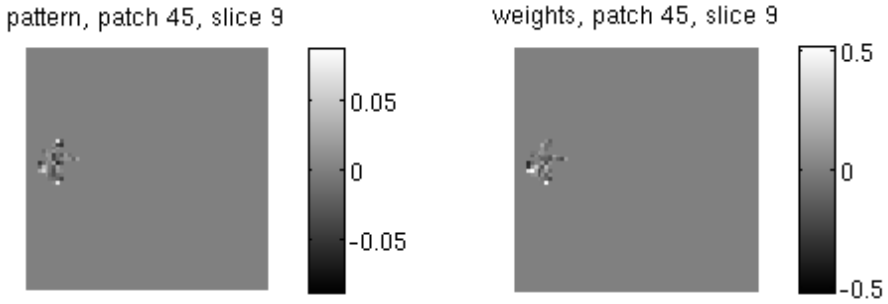

Figure 25: Comparison of patterns and of subject 7 in V1 (slice 9) [2].

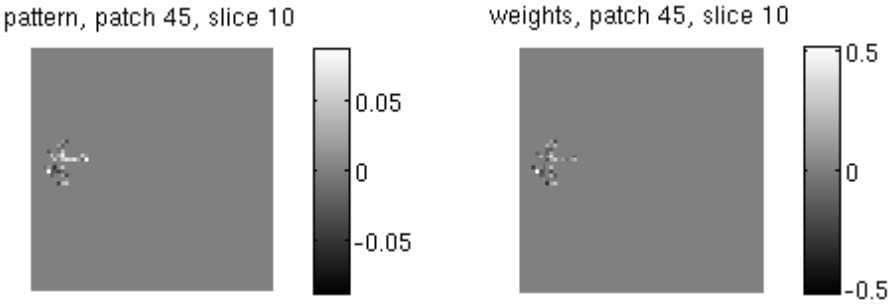

Figure 26: Comparison of patterns and weights of subject 7 in V1 (slice 10) [2].

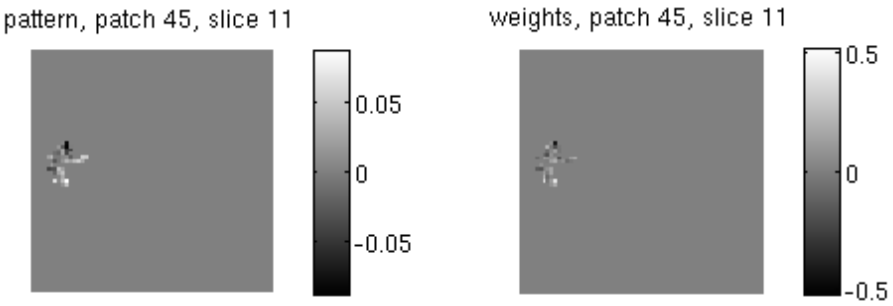

Figure 27: Comparison of patterns and weights of subject 7 in V1 (slice 11) [2].

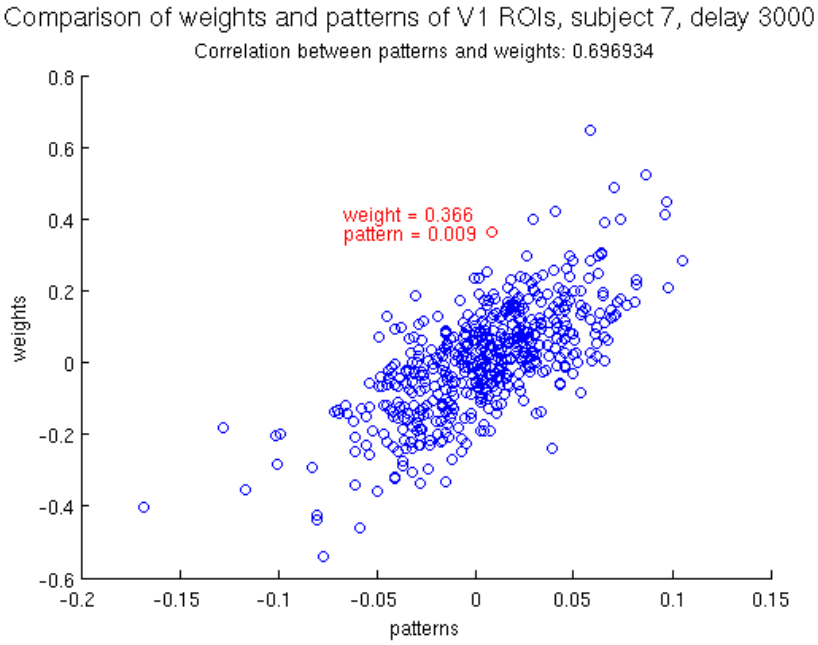

Figure 28: Overall comparison of patterns and weights of V1 voxels of subject 7 [2]. Marked in red (with respective weight and pattern written next to it in red) is one example in which there is a big discrepancy between pattern and weight value.