Gaussian processes for whole-brain feature selection and classification in fMRI

Ekaterina I Lomakina[1,2] · Kay H Brodersen[1,2] · Timothy E J Behrens[3] · Klaas E Stephan[2,4] · Joachim M Buhmann[1]

[1] Department of Computer Science, ETH Zurich, Switzerland [2] Laboratory for Social and Neural Systems Research, Department of Economics, University of Zurich, Switzerland [3] Centre for Functional Magnetic Resonance Imaging of the Brain (FMRIB), John Radcliffe Hospital, University of Oxford, United Kingdom [4] Wellcome Trust Centre for Neuroimaging, University College London, United Kingdom

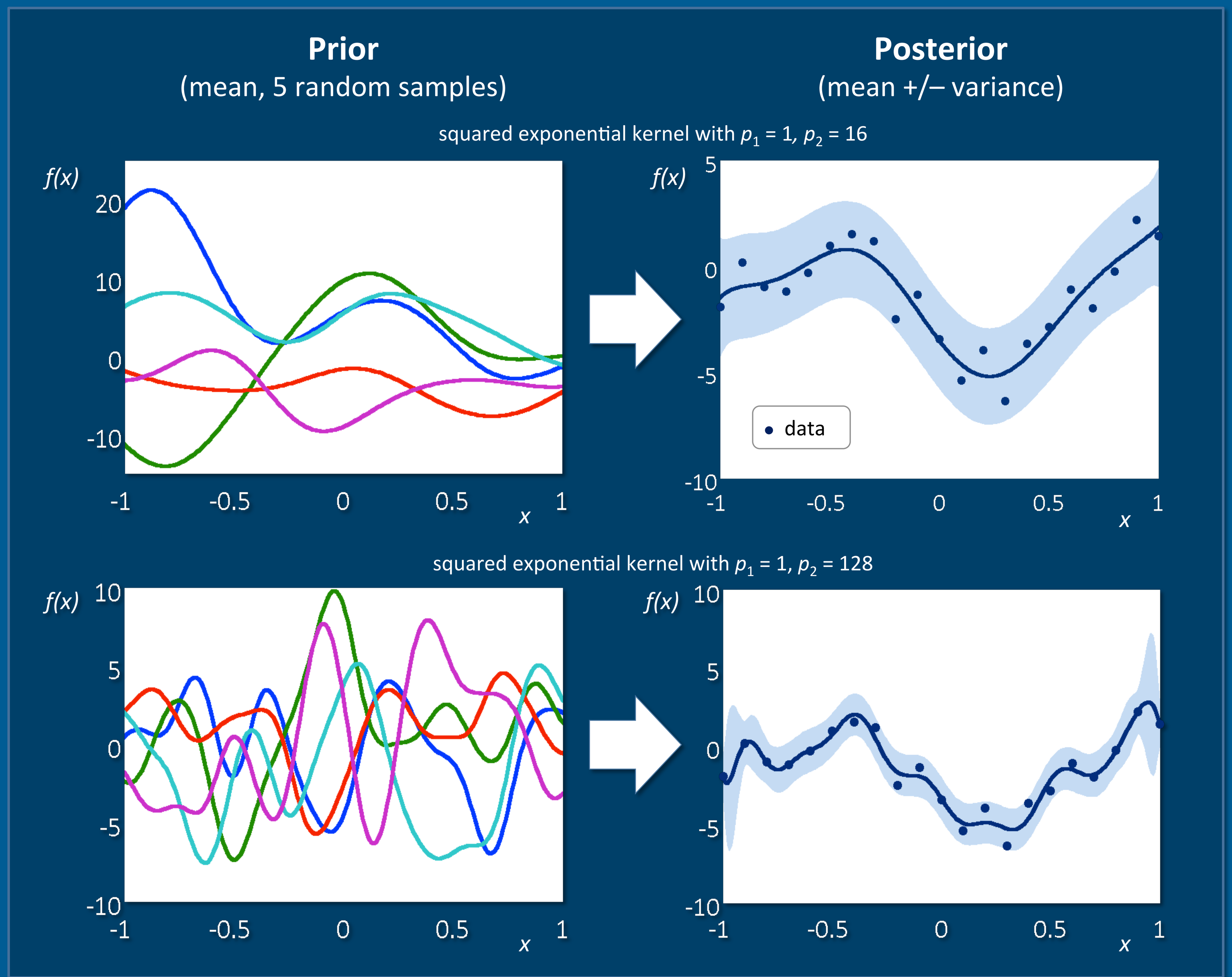
1 Summary

- Multivariate decoding approaches for fMRI attempt to infer cognitive or perceptual states from distributed measurements of brain activity. While support vector machines (SVM) are routinely used for multivariate fMRI investigations, other methods, such as Gaussian process models, may offer advantages in dealing with whole-brain data.
- The critical challenge for all decoding models is the high dimensionality of fMRI data. We address this challenge by combining Gaussian processes with parametric permutation tests on voxel weights. This approach allows us to (i) study the spatial deployment of jointly informative voxels, and (ii) select informative features for a whole-brain decoding model.
- We illustrate the utility of our approach by analysing fMRI data from a simple decision-making task.

2 Gaussian processes for fMRI

Models that predict brain states from whole-brain activity must deal with the high dimensionality of fMRI data by regularizing the inference problem.

Gaussian processes (GP) represent a powerful nonparametric Bayesian approach that addresses this requirement using a prior over mapping functions [1,2].

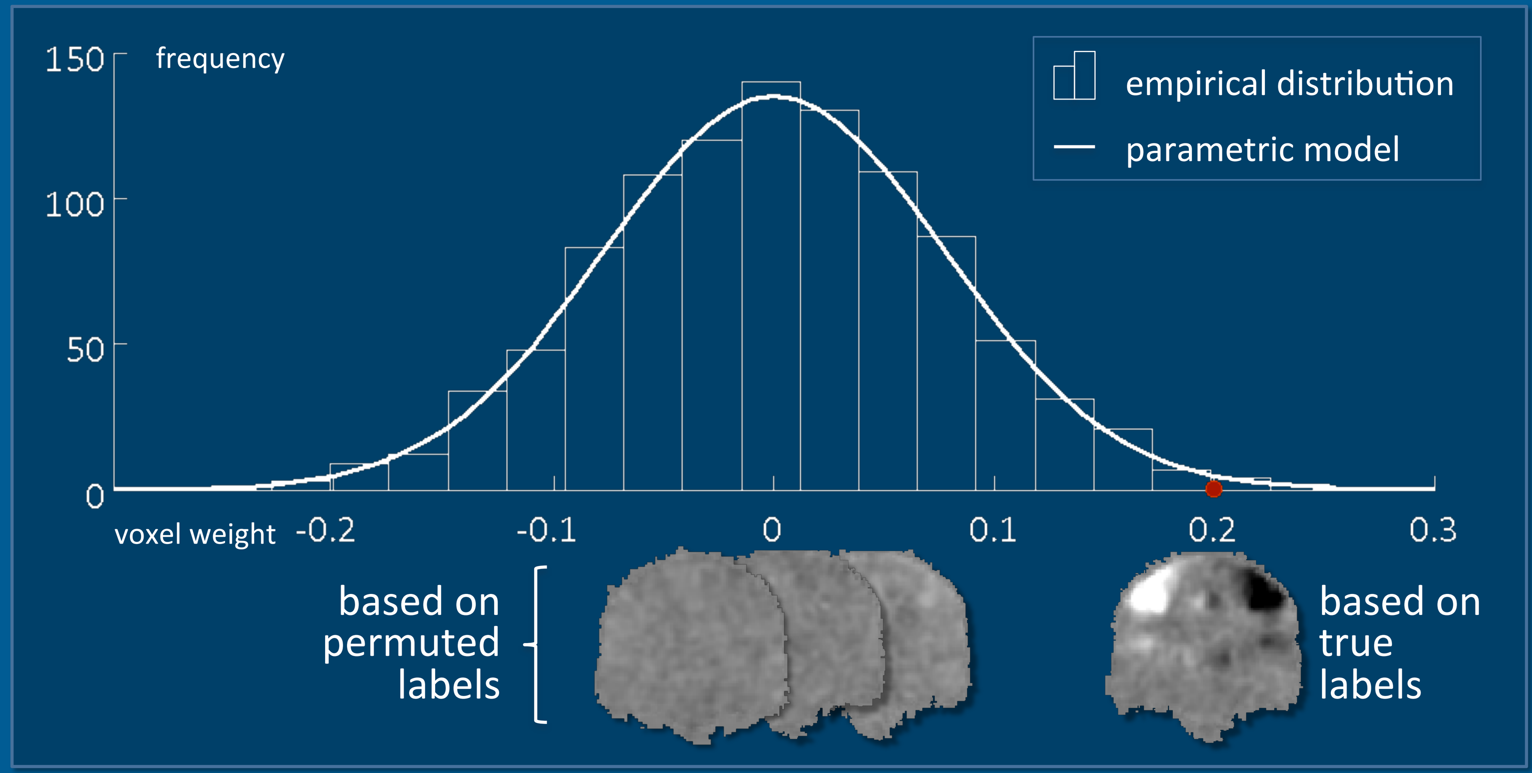


In this study, we design a GP model for reconstructing the spatial deployment of voxels that are jointly informative of a cognitive state of interest.

3 Statistical inference on voxel weights

When using Gaussian processes by themselves, the estimated role of individual voxelы may heavily depend on task-unrelated properties of the data, e.g., signal variability over time.

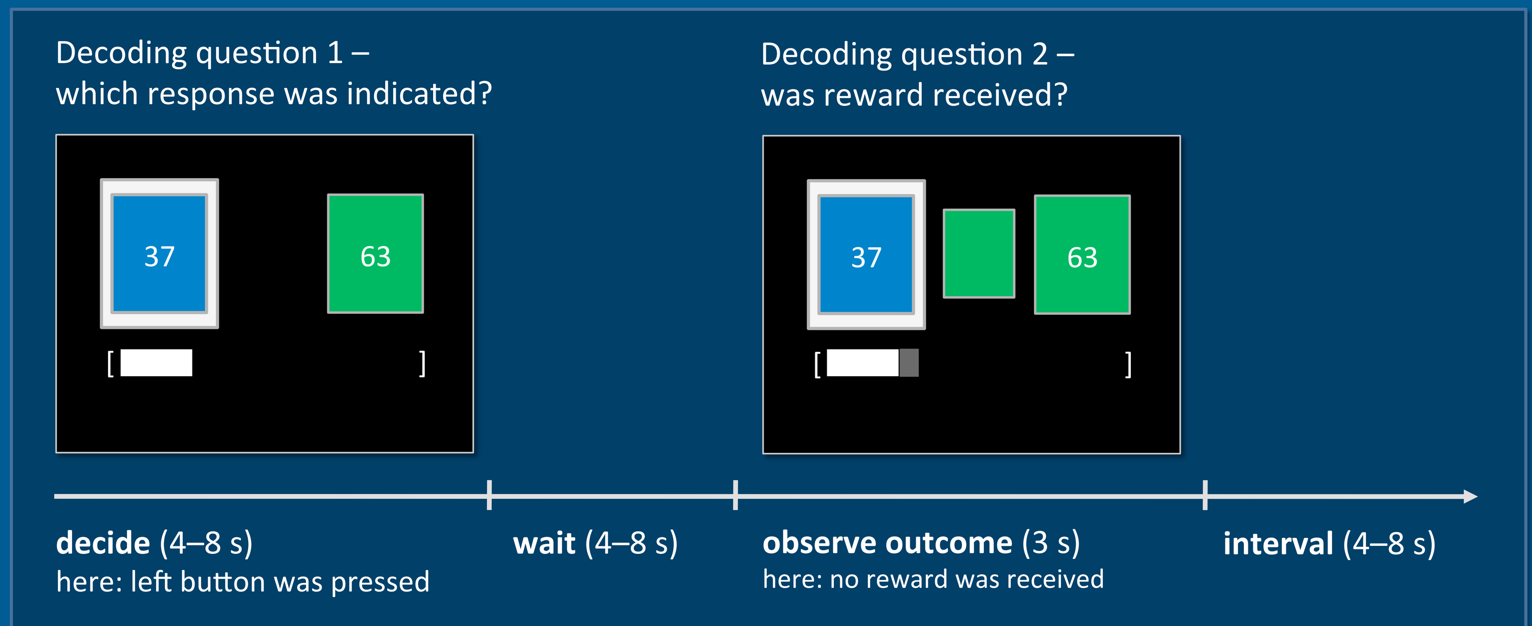
We address this issue by estimating the empirical null distribution of voxel weights. This distribution can be obtained by repeatedly re-estimating the model based on randomly permuted labels.



The permutation test allows us to compute a map of t-scores that express how significantly each voxel contributed to a distributed encoding of a cognitive state.

4 Experimental design

We illustrate our approach using an fMRI dataset (n = 16) acquired in the context of a simple decision-making task [3]. On each trial, participants were asked to choose between a blue and a green option using a button press (left/right index finger). Rewards were associated with the two cards according to probabilities that varied over time. Potential rewards were displayed on the cards.



Using a linear SVM, we decoded, on a trial-by-trial basis, which response was given and whether a reward was received.

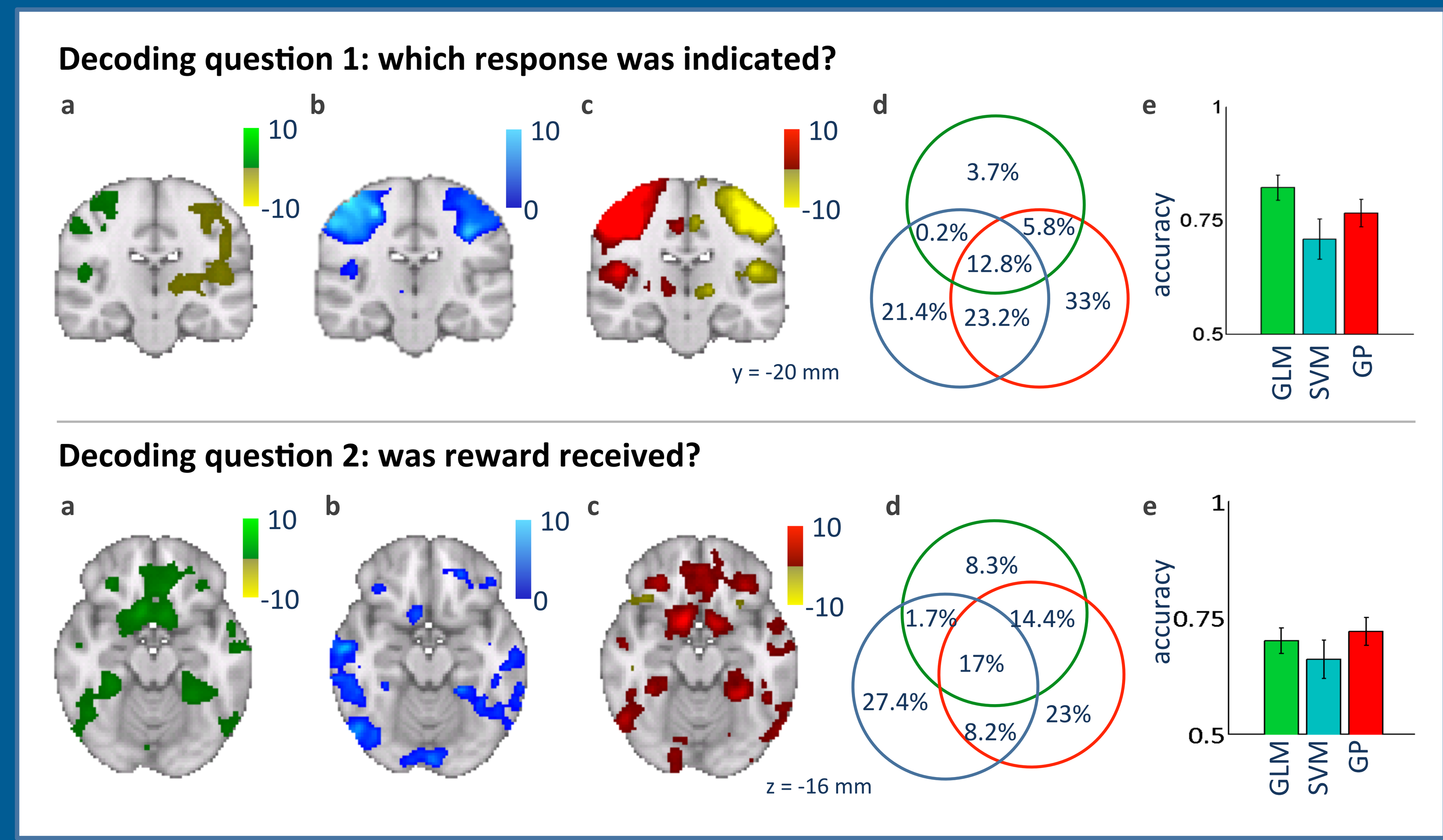
5 Results

We constructed discriminative maps using three competing methods:

(a) a mass-univariate model using a conventional GLM
(b) a locally multivariate searchlight approach using a linear SVM
(c) the proposed whole-brain analysis using Gaussian processes (GP)

When thresholding all three maps at the same level of specificity (p = 0.001 unc.), Gaussian process maps displayed the highest sensitivity (d).

We then used these methods for fold-wise feature selection in a linear SVM (e). In this instance, Gaussian processes afforded no significant increase over conventional methods although classification accuracies remained competitive.



6 Conclusions

- We have proposed a combination of Gaussian processes and parametric permutation tests. Our approach aims to address the challenge of high-dimensionality of fMRI data without assuming specific spatial response features.
- Our approach (i) can be used for constructing a discriminative map, (ii) for feature selection in classification, and (iii) it is computationally efficient (<1 h per subject).
- Using an example from decision making, our approach achieves greater sensitivity than both mass-univariate and locally multivariate searchlight models.

Acknolwedgements
The research was supported by the UK Medical Research Council (TB).

References
1. Marquand, A. et al., 2010. Quantitative prediction of subjective pain intensity from whole-brain fMRI data using Gaussian processes. *NeuroImage*, 49(3), 2178-2189.
2. Rasmussen, C.E. & Williams, C.K.I., 2006. Gaussian Processes for Machine Learning, MIT Press.
3. Behrens, T.E.J. et al., 2007. Learning the value of information in an uncertain world. *Nature Neuroscience*, 10, 1214-1221.
4. Haynes, J. & Rees, G., 2006. Decoding mental states from brain activity in humans. *Nature Reviews Neuroscience*, 7(7), 523-534.
5. Kriegeskorte, N. et al., 2006. Information-based functional brain mapping. *PNAS*, 103(10), 3863-3868.