# Variational Bayesian mixed-effects inference for classification studies


ETH Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

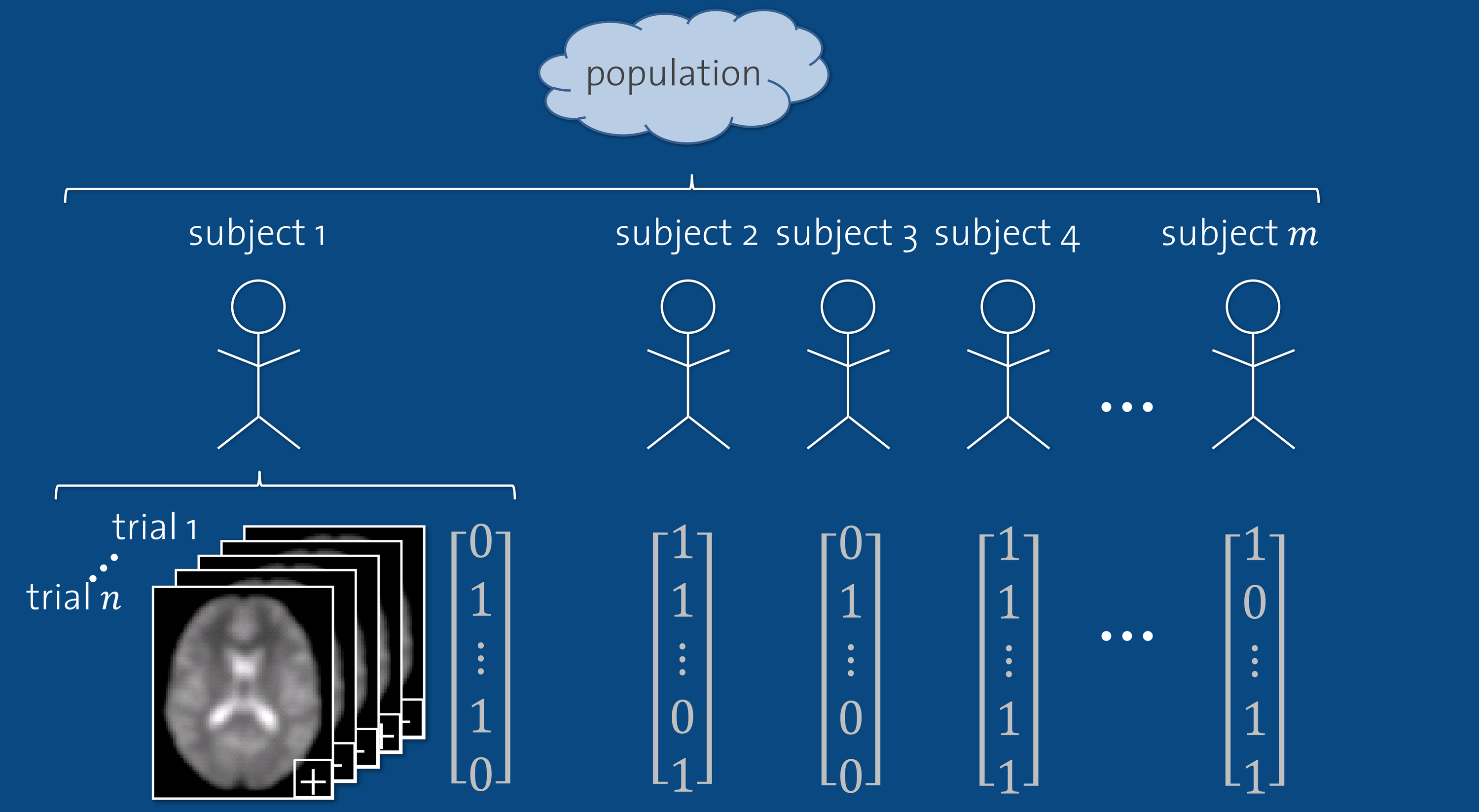
UCL

Kay H Brodersen[1,2,3] · Jean Daunizeau[4] · Christoph Mathys[1,3] · Justin R Chumbley[3] · Joachim M Buhmann[2] · Klaas E Stephan[1,3,5]

[1] Translational Neuromodeling Unit (TNU), University of Zurich & ETH Zurich, Switzerland [2] Department of Computer Science, ETH Zurich, Switzerland [3] Laboratory for Social and Neural Systems Research, Department of Economics, University of Zurich, Switzerland [4] Brain and Spine Institute, Paris, France [5] Wellcome Trust Centre for Neuroimaging, University College London, United Kingdom


## 1 Introduction

Multivariate classification algorithms are powerful tools for predicting cognitive or pathophysiological states from neuroimaging data [1]. However, there are unresolved questions regarding the statistical evaluation of classification performance in group studies.

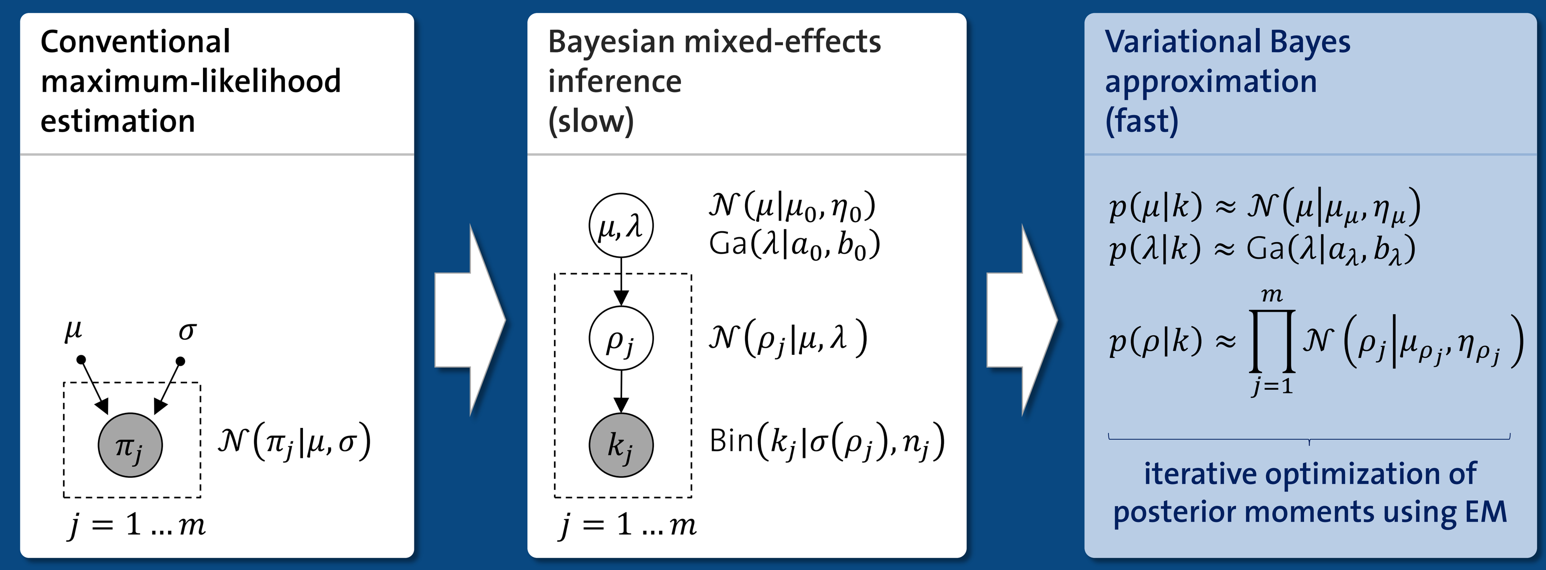



Inference in group studies requires models that explicitly account for fixed-effects (within-subjects) and random-effects (across-subjects) variance components. This principle is standard practice in mass-univariate analyses [2–5] but has not yet been adopted in classification studies.

## 2 Variational mixed-effects inference

We recently introduced Bayesian models for mixed-effects inference in multivariate classification studies [6,7]. Until now, their practical utility was limited by the high computational complexity of the underlying Markov chain Monte Carlo (MCMC) sampling algorithms.

Here, we solve this remaining problem by introducing a variational Bayes (VB) approach to inference [8].




## 3 Model assumptions and algorithm

$$p(\mu,\lambda,\rho|k)$$

variational inference

$$p(\mu,\lambda,\rho|k) \approx q(\mu,\lambda,\rho) \propto \exp\big(I(\mu,\lambda,\rho)\big)$$
$$I(\mu,\lambda,\rho) := \ln p(k,\mu,\lambda,\rho)$$

mean-field approximation

$$q(\mu,\lambda,\rho) = q(\mu)q(\lambda)q(\rho)$$

parametric assumptions

$$q(\mu) = \mathcal{N}(\mu|\mu_\mu,\eta_\mu), \quad q(\lambda) = \mathrm{Ga}(\lambda|a_\lambda,b_\lambda), \quad q(\rho) = \textstyle\prod_{j=1}^{m} \mathcal{N}\left(\rho_j\middle|\mu_{\rho_j},\eta_{\rho_j}\right)$$

expectation maximization

$$I_1(\mu) = \langle \ln p(k,\rho,\mu,\lambda)\rangle_{q(\rho,\lambda)}$$
$$\Rightarrow \mu_\mu = \frac{\mu_0\eta_0 + a_\lambda b_\lambda \sum_{j=1}^{m}\mu_{\rho_j}}{\eta_0 + m\,a_\lambda b_\lambda}, \quad \eta_\mu = m\,a_\lambda b_\lambda + \eta_0$$

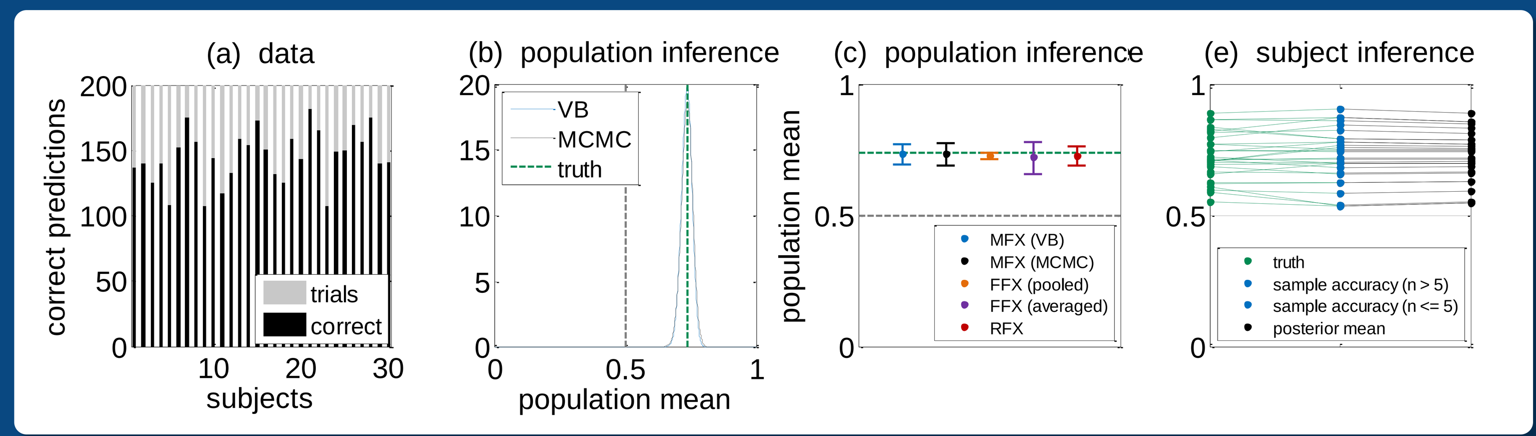
until convergence

$$I_2(\lambda) = \langle \ln p(k,\rho,\mu,\lambda)\rangle_{q(\rho,\mu)}$$
$$\Rightarrow a_\lambda = \frac{1}{2}m + a_0$$
$$b_\lambda = \left(\frac{1}{b_0} + \frac{1}{2}\sum_{j=1}^{m}\left(\left(\mu_{\rho_j} - \mu_\mu\right)^2 + \eta_{\rho_j}^{-1} + \eta_\mu^{-1}\right)\right)^{-1}$$

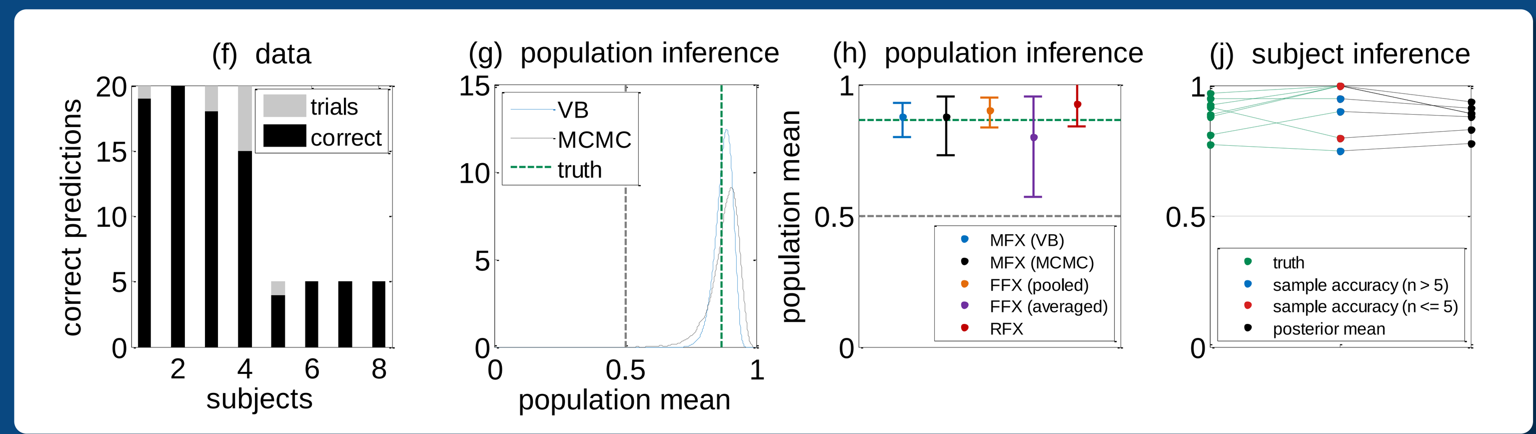
$$I_3(\rho) = \langle \ln p(k,\rho,\mu,\lambda)\rangle_{q(\mu,\lambda)}$$
$$\rho^* \leftarrow \rho^* - \left[\left.\frac{d^2 I(\rho)}{d\rho^2}\right|_{\rho=\rho^*}\right]^{-1} \times \left.\frac{dI(\rho)}{d\rho}\right|_{\rho=\rho^*}$$
$$\mu_\rho = \rho^*, \quad \eta_\rho = \mathrm{Diag}\left(-\left.\frac{d^2 I(\rho)}{d\rho^2}\right|_{\rho=\rho^*}\right)$$

## 4 Illustrative examples

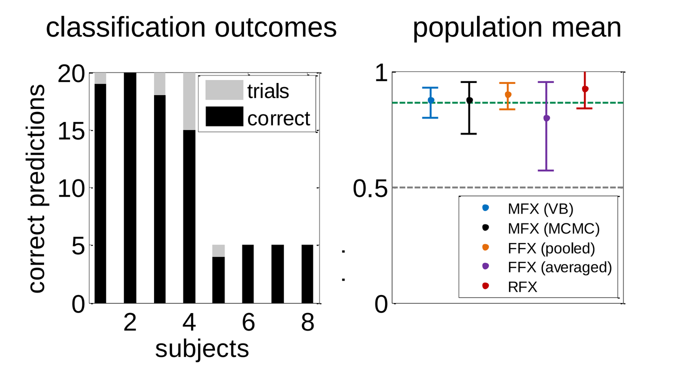
In a first setting, we generated synthetic classification outcomes for a group of 20 subjects with 200 trials each. VB inferences were practically indistinguishable from an MCMC reference implementation. Fixed-effects approaches, by contrast, were over- or underconfident.




In a second simulation, we considered a smaller, more heteroscedastic group. VB and MCMC posterior means agreed nicely. Fixed-effects methods were over- or underconfident, and a conventional random-effects t-test provided an uninterpretable confidence interval that included accuracies above 100%.

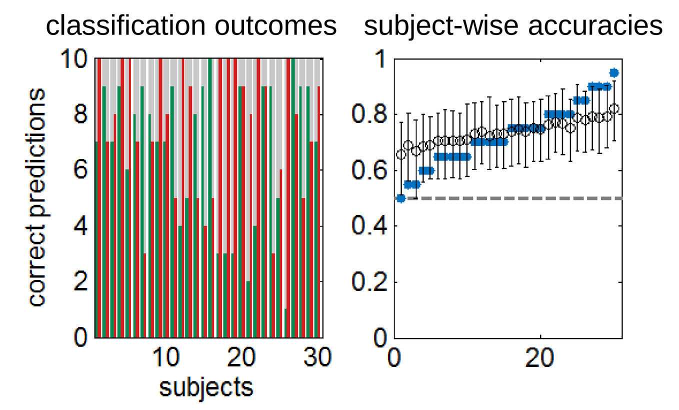



## 5 Advantages over conventional methods



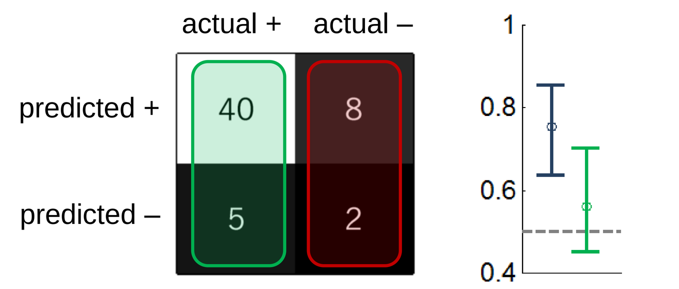

**1 | Inference on the population mean**

In contrast to the proposed Bayesian posterior interval (blue), the conventional confidence interval (red) is biased, and a conventional fixed-effects analysis (orange) is overconfident.



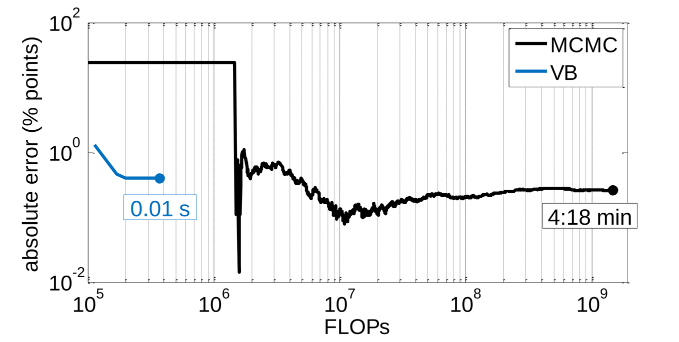

**2 | Inference on individual accuracies**

Unlike conventional sample accuracies (blue), the proposed Bayesian posterior means of individual accuracies (black) are informed by data from (or 'shrunk towards') the entire group. This prevents overfitting and yields more precise estimates.




**3 | Inference on the balanced accuracy**

The balanced accuracy (green) provides a more useful measure of classification performance than the accuracy (blue), especially in the context of imbalanced data. The example illustrates this in a single-subject setting.




**4 | Computational efficiency**

While MCMC ultimately achieves a marginally lower error, VB is computationally more efficient by approximately 4 orders of magnitude.

## 6 Conclusions

- Bayesian mixed-effects inference for group studies (i) models within-subject and across-subjects uncertainty, (ii) enables conclusions in terms of intuitive probability statements, and (iii) can be used with various performance measures, e.g., the balanced accuracy. In contrast to previous sampling algorithms, our variational approximation is computationally highly efficient.
- We hope that our approach will improve the precision and interpretability of statistical inference in future classification studies.


**References**

1. Haynes, J., & Rees, G., 2006. Decoding mental states from brain activity in humans. *Nature Reviews Neuroscience*, 7(7), 523-534.
2. Holmes, A. P., & Friston, K. J. (1998). Generalisability, random effects and population inference. Fourth Int. Conf. on Functional Mapping of the Human Brain, *NeuroImage*, 7, S754.
3. Beckmann, C. F., Jenkinson, M., & Smith, S. M. (2003). General multilevel linear modeling for group analysis in FMRI. *NeuroImage*, 20(2), 1052-1063.
4. Woolrich, M. W., Behrens, T. E. J., Beckmann, C. F., Jenkinson, M., & Smith, S. M. (2004). Multilevel linear modelling for FMRI group analysis using Bayesian inference. *NeuroImage*, 21(4), 1732–1747.
5. Mumford, J. A., & Nichols, T. (2009). Simple group fMRI modeling and inference. *NeuroImage*, 47(4), 1469-1475.
6. Brodersen, K.H., Ong, C.S., Buhmann, J.M., & Stephan, K.E. (2010). The balanced accuracy and its posterior distribution. *ICPR*, 3121-3124.
7. Brodersen, K.H., Mathys, C., Chumbley, J.R., Daunizeau, J., Ong, C.S., Buhmann, J.M., & Stephan, K.E. (*in revision*) Mixed-effects inferene on classification performance in hierarchical datasets.
8. Brodersen, K.H., Daunizeau, J., Mathys, C., Chumbley, J.R., Buhmann, J.M., & Stephan, K.E. (*in preparation*) Variational Bayesian mixed-effects inference for classification studies.


MATLAB code available online