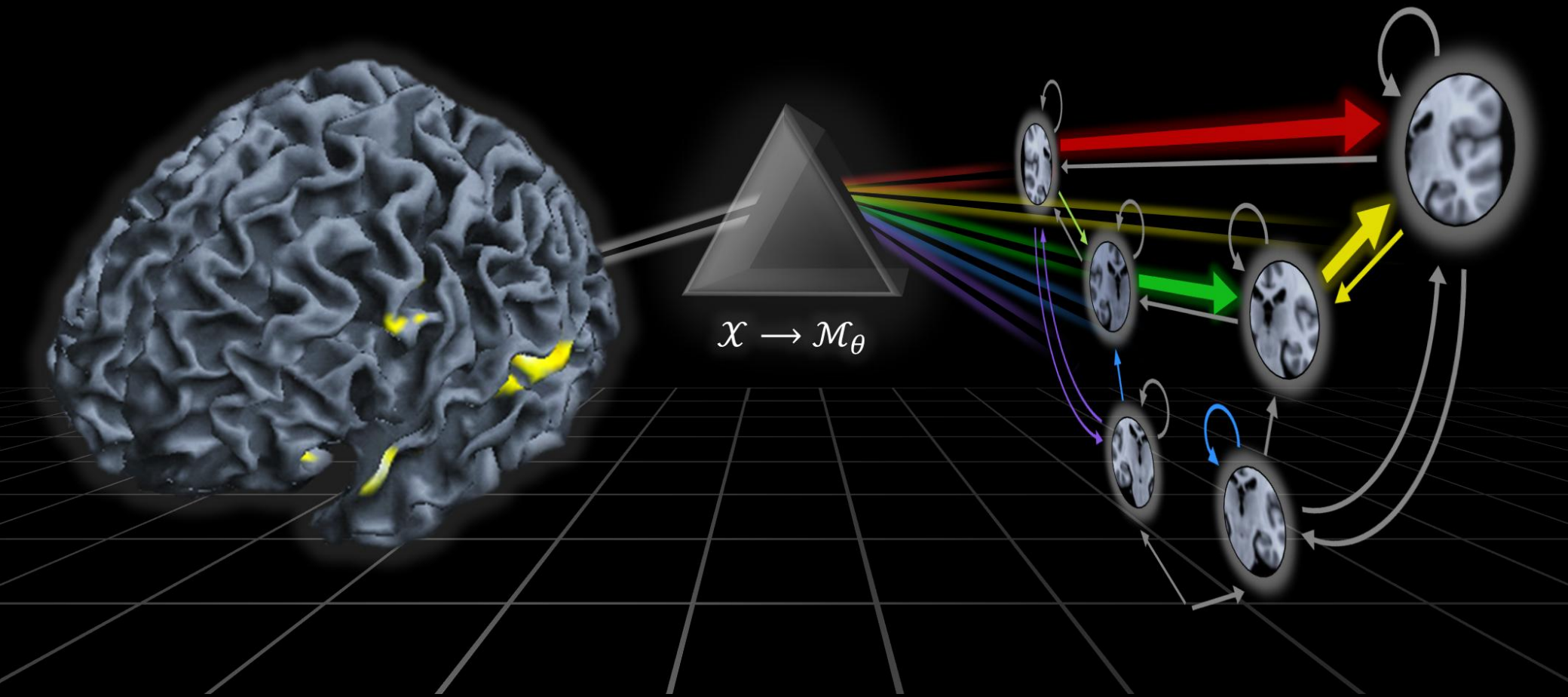
Generative embedding and variational Bayesian inference for multivariate time series

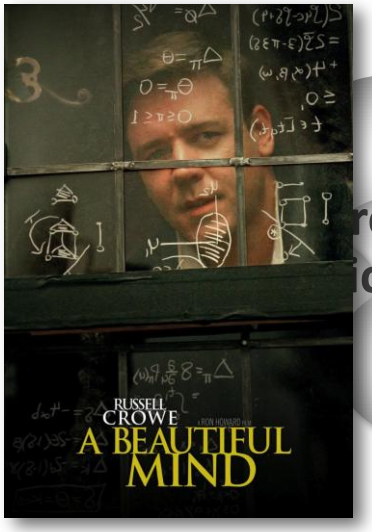
Kay H. Brodersen[1,2]

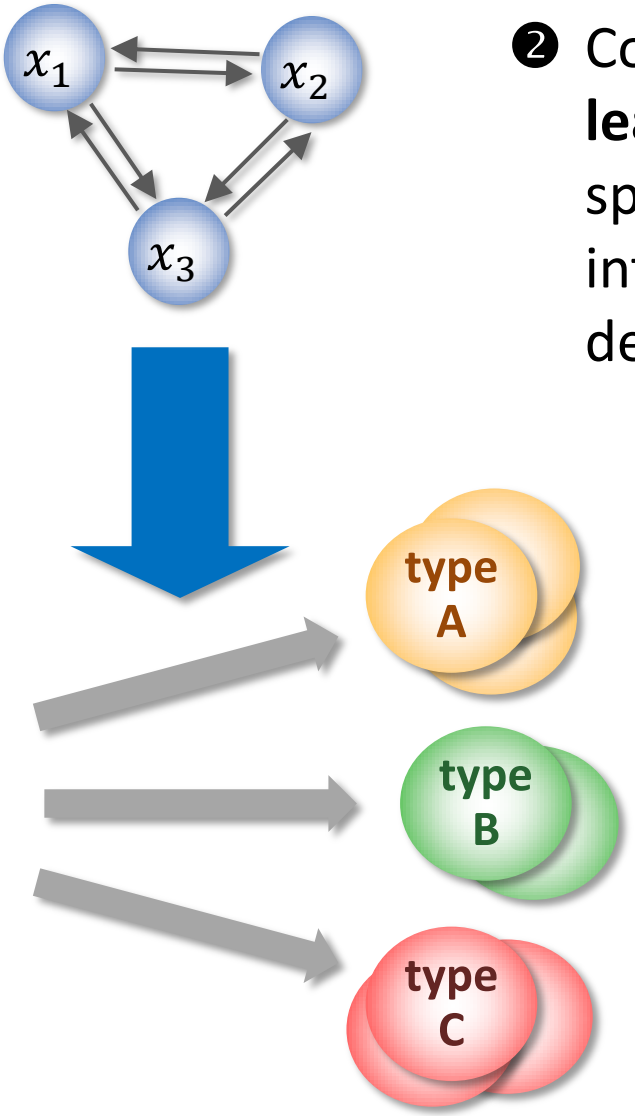
[1] Machine Learning Laboratory, Department of Computer Science, ETH Zurich
[2] Translational Neuromodeling Unit (TNU), Institute of Biomedical Engineering, University of Zurich & ETH Zurich



A computational approach to dissecting spectrum disorders

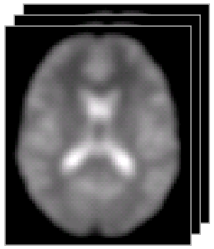
❶ **Psychiatry** lacks pathophysiologically informed diagnostic classifications.





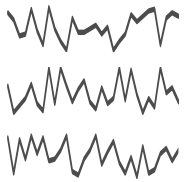
❷ Could **machine learning** help dissect spectrum disorders into mechanistically defined subgroups?

Model-based analysis by generative embedding



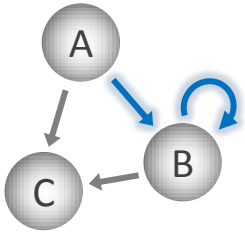
measurements from an individual subject

step 1 — extraction



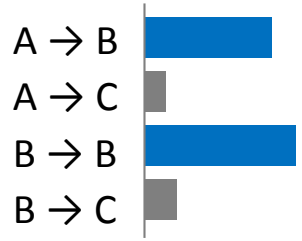
time series in regions of interest

step 2 — modelling



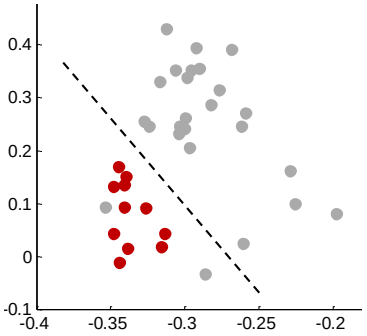
subject-specific generative model

step 3 — embedding



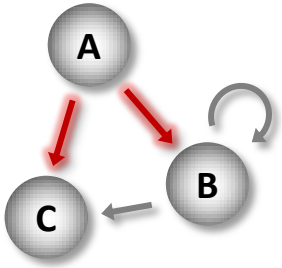
representation in model-based feature space

step 4 — analysis



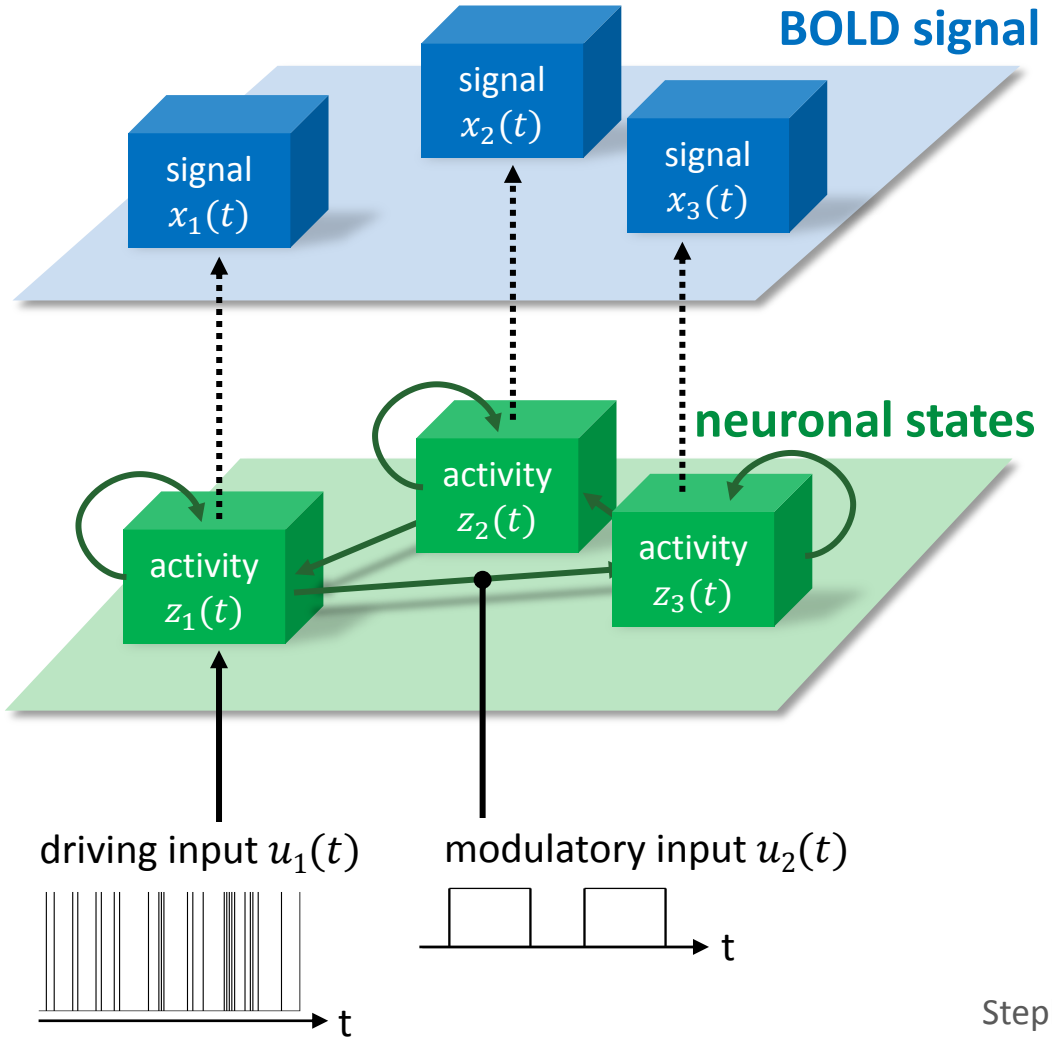
model of group structure

step 5 — interpretation



jointly discriminative connection strengths

Choosing a generative model: DCM for fMRI



haemodynamic forward model

$$x = g(z, \theta_h)$$

neural state equation

$$\dot{z} = \left(A + \sum u_j B^{(j)}\right) z + Cu$$

intrinsic connectivity

modulation of connectivity

direct inputs

Friston, Harrison & Penny (2003) *NeuroImage*
Stephan & Friston (2007) *Handbook of Brain Connectivity*

Constructing a classifier

Y_t → f → that k which maximizes $p(X_t = k|Y_t, X, Y)$

Generative classifiers

use Bayes' rule to obtain $p(X_t|Y_t) \propto p(Y_t|X_t)p(X_t)$

- *Gaussian Naïve Bayes*
- *Linear Discriminant Analysis*
- *Gaussian processes*

Discriminative classifiers

estimate $p(X_t|Y_t)$ directly without Bayes' theorem

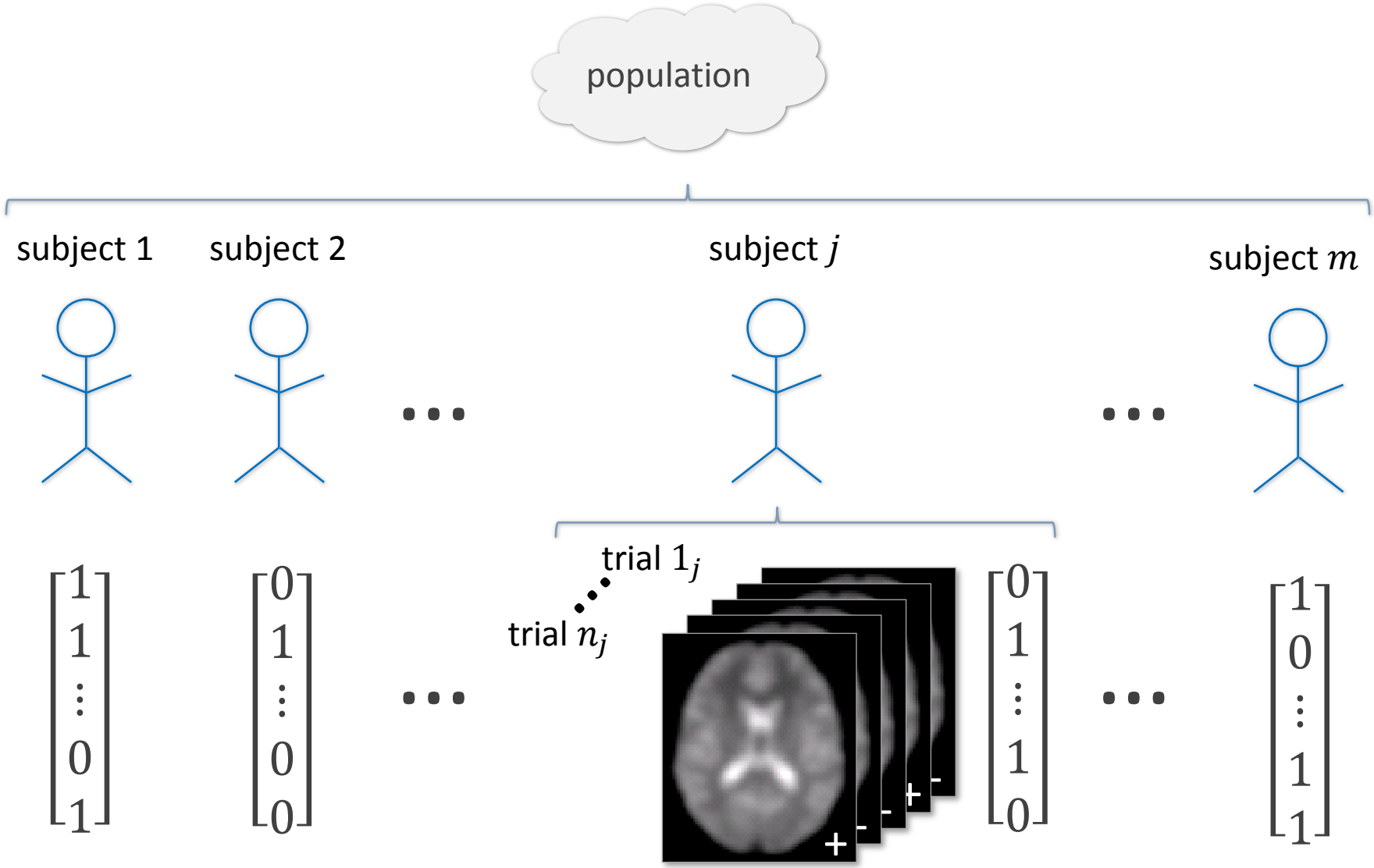
- *Logistic regression*
- *Relevance Vector Machine*

Discriminant classifiers

estimate $f(Y_t)$ directly

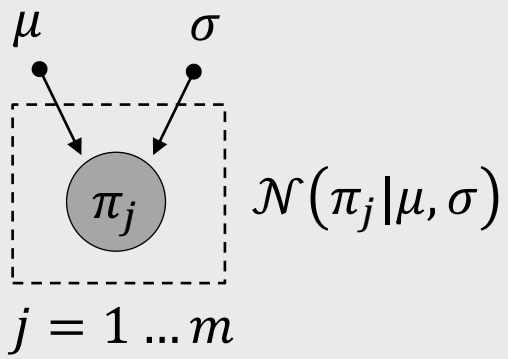
- *Fisher's Linear Discriminant*
- *Support Vector Machine*

Hierarchical classification analyses

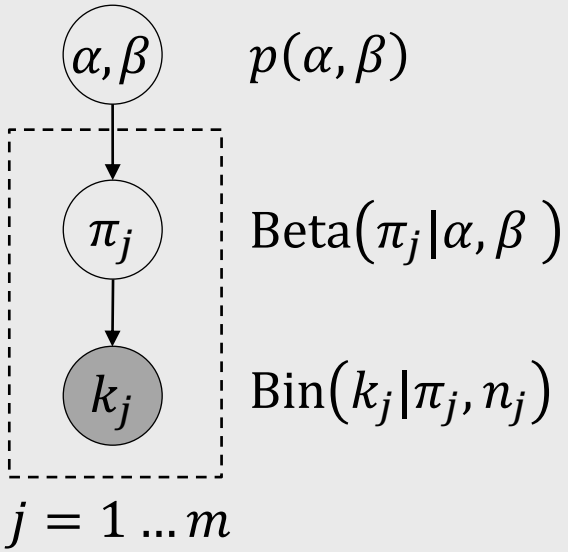


Towards a mixed-effects model for performance evaluation

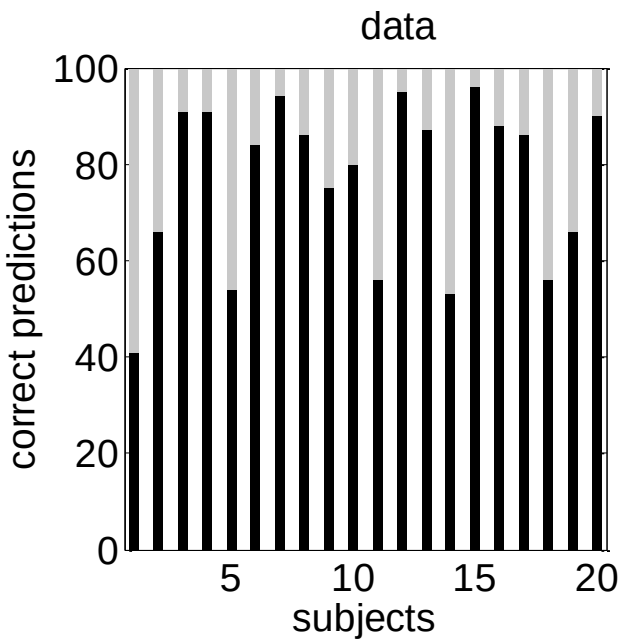
Maximum-likelihood random-effects estimation

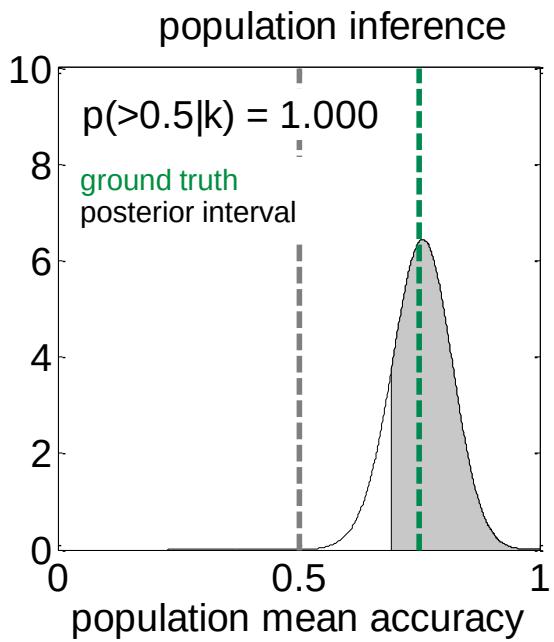


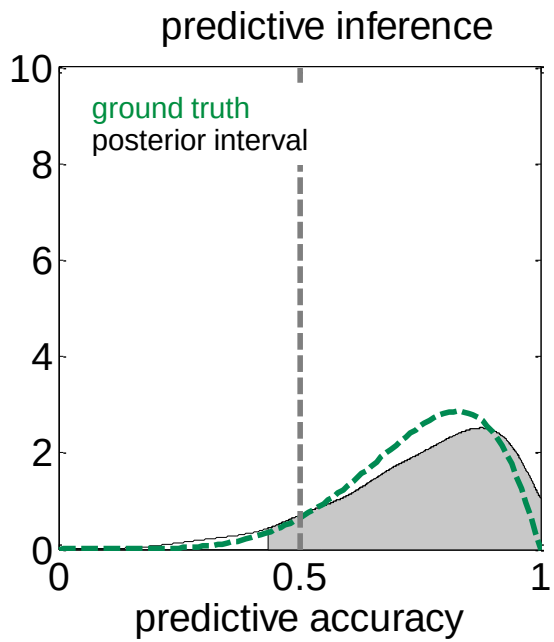
Bayesian mixed-effects inference



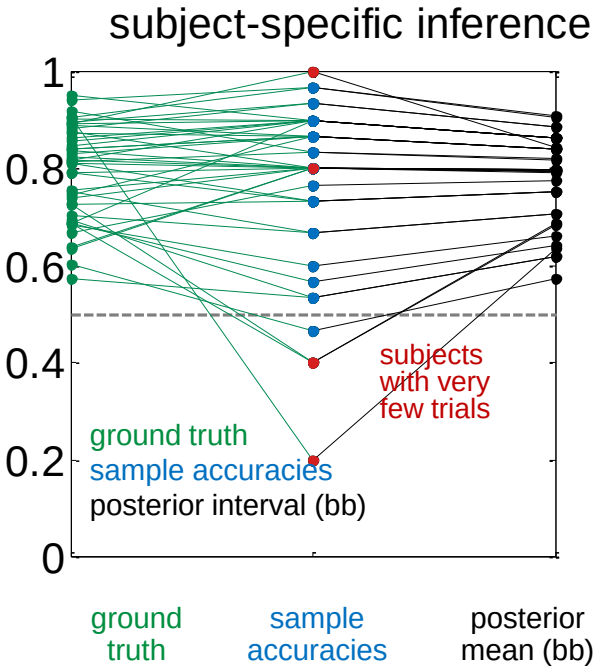
Population inference on synthetic classification outcomes

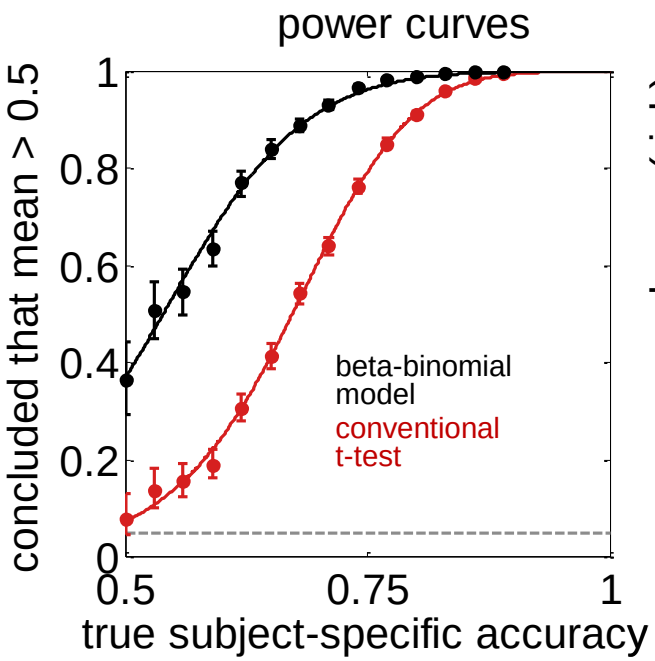


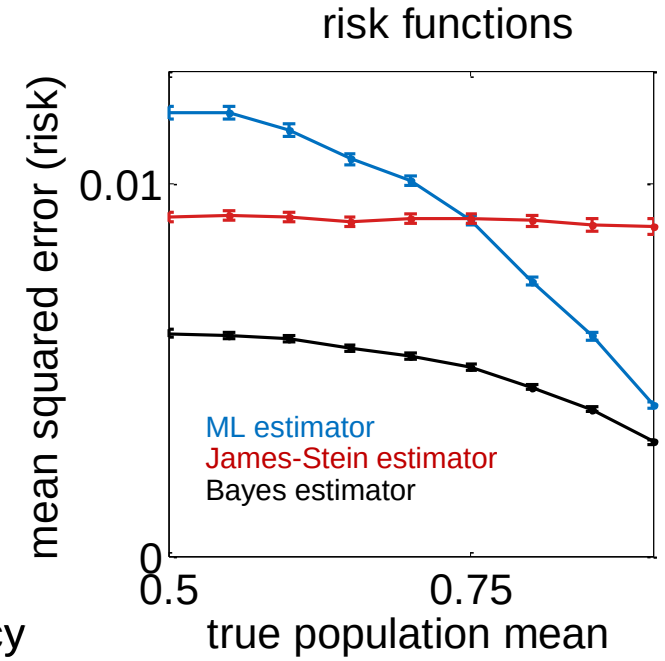




Subject-specific inference on synthetic classification outcomes

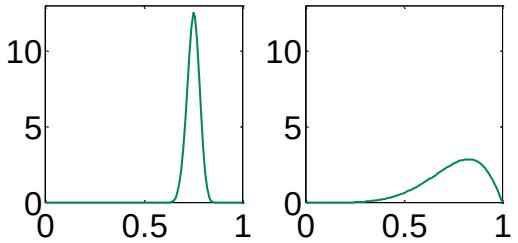


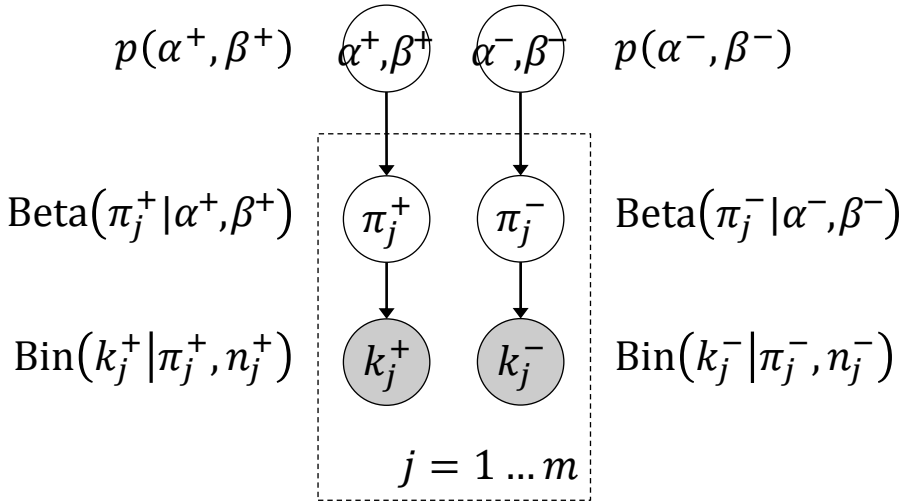




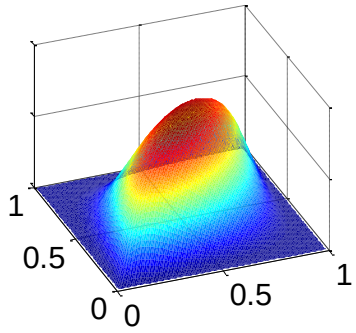
Mixed-effects inference on the balanced accuracy

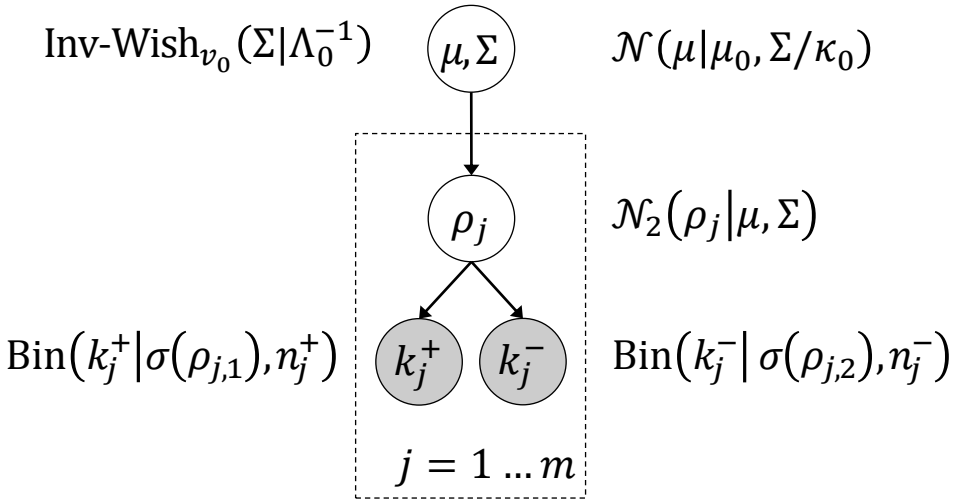
Beta-binomial model



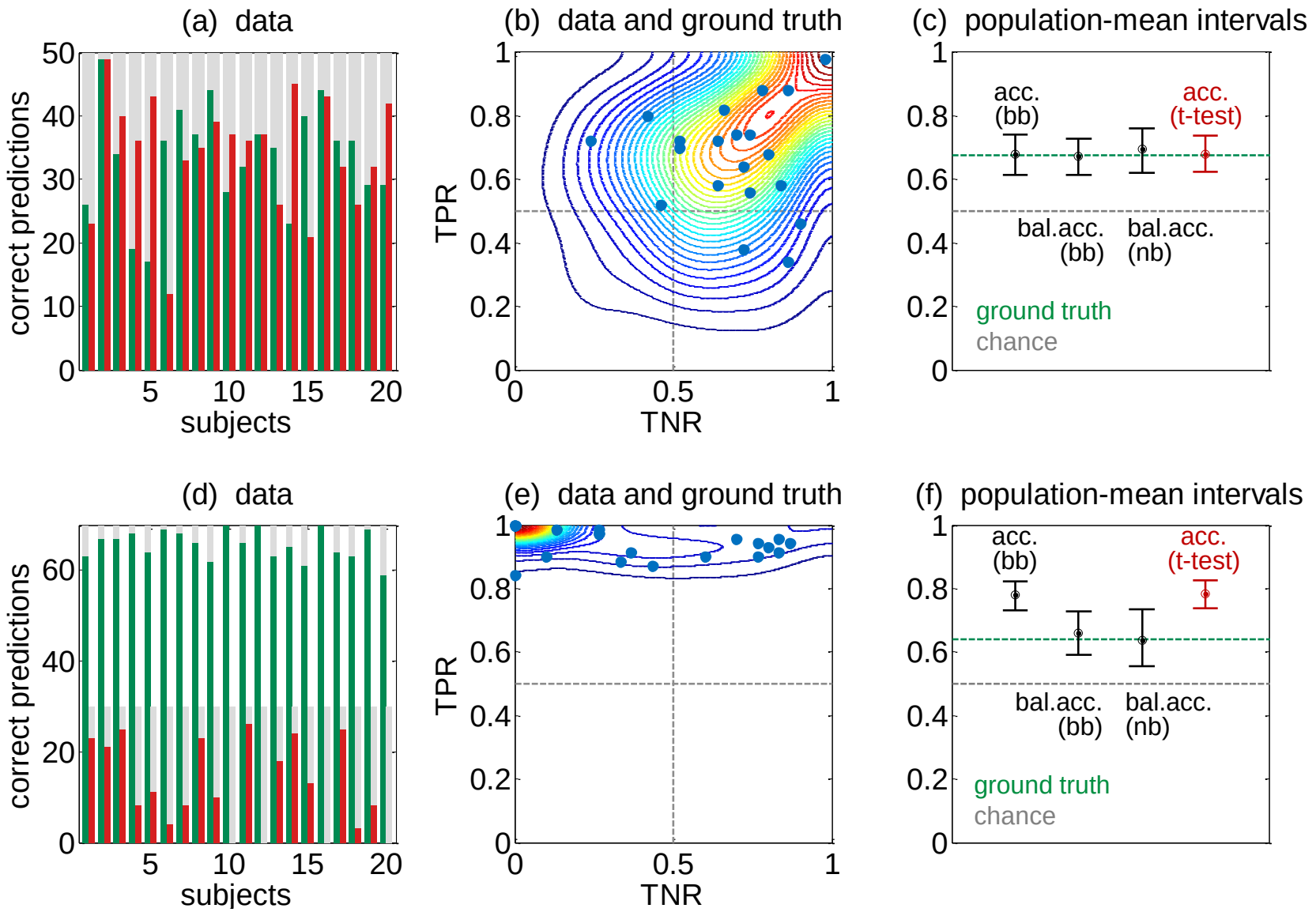


Bivariate normal-binomial model



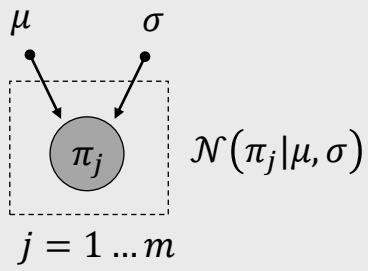


Application to synthetic data features

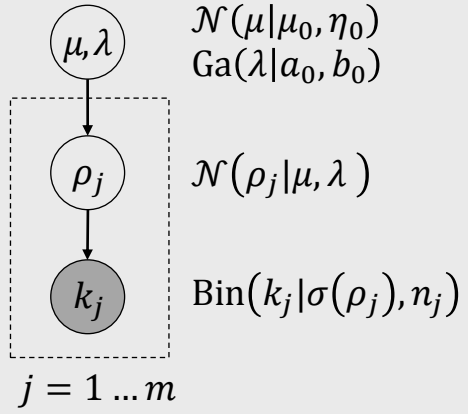


Towards a variational approach

Conventional maximum-likelihood estimation



Bayesian mixed-effects inference



Variational Bayes approximation

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

iterative optimization of posterior moments

Variational algorithm

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

variational inference

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

mean-field approximation

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

parametric assumptions

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

variational algorithm with Laplace approximations

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

until convergence

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

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

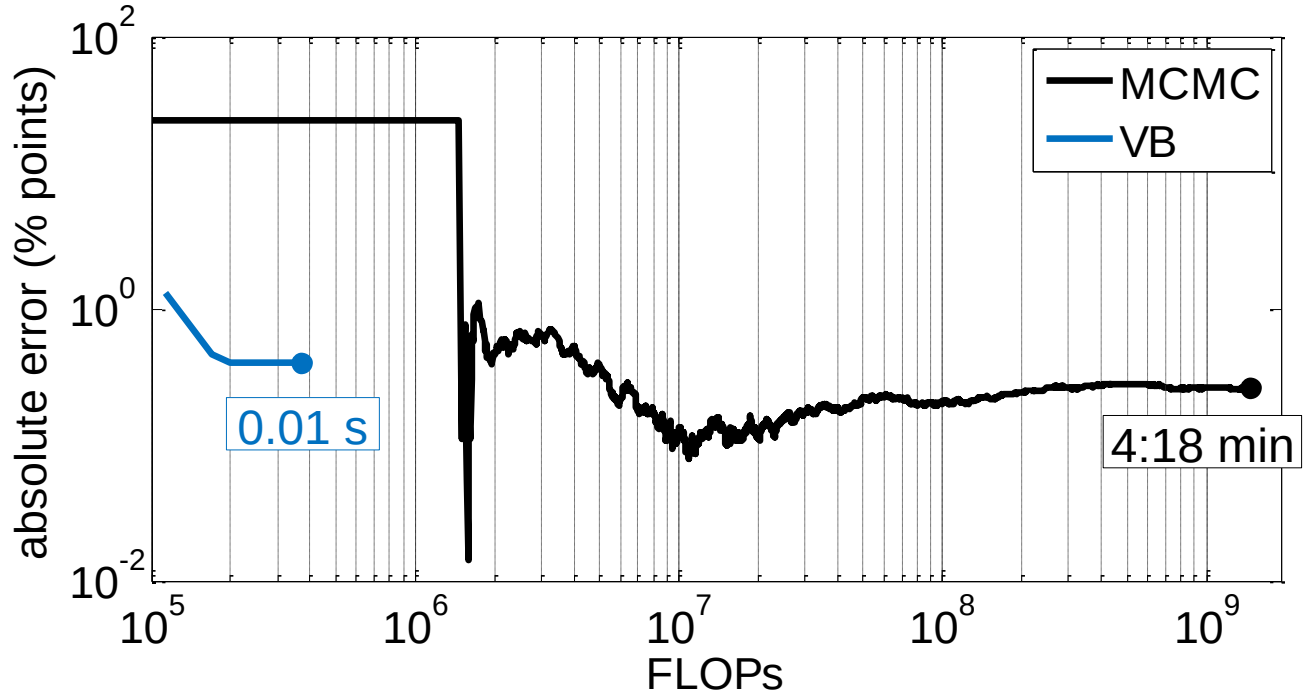
Computational complexity

Example

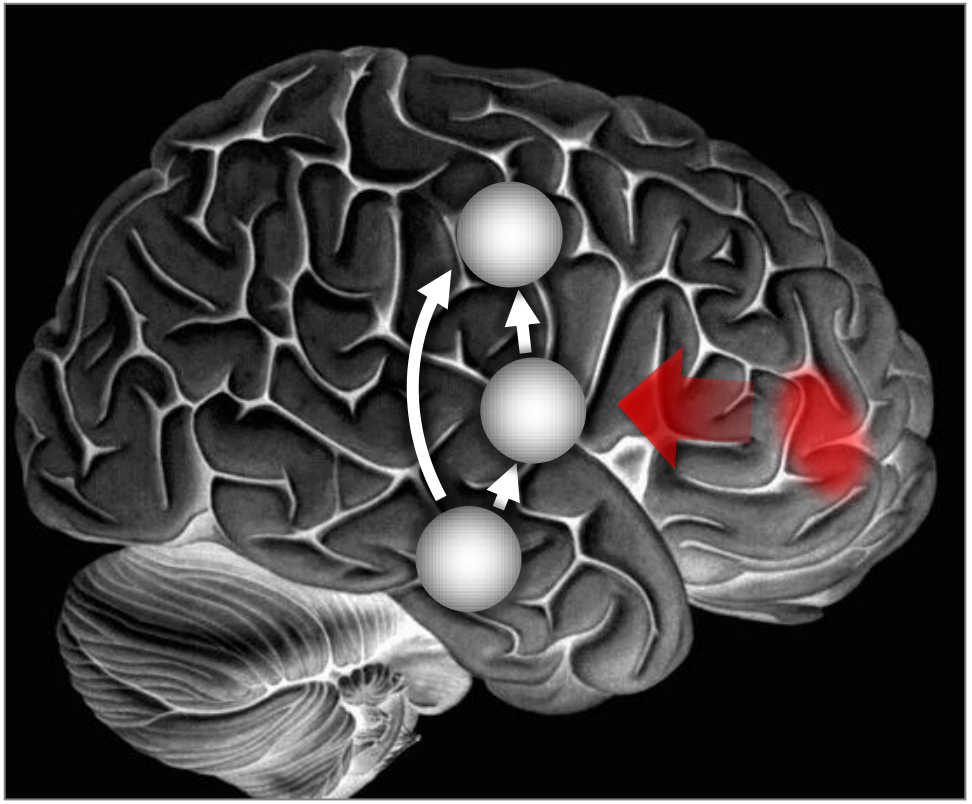
whole-brain (50,000 voxels)
mass-univariate evaluation of classification accuracy
20 subjects

Variational Bayes
0.009 s/voxel
7:30 min

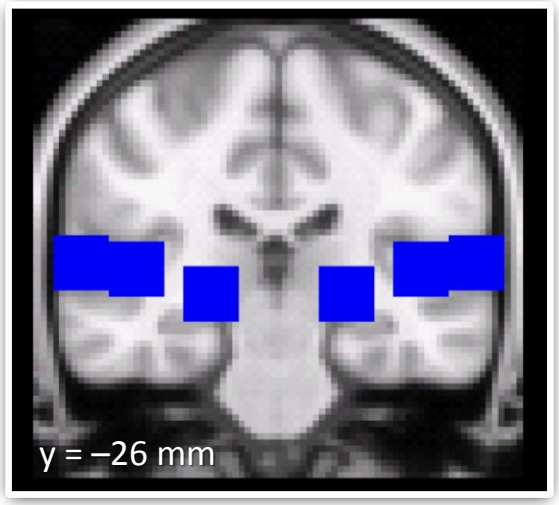
Estimation error & computational complexity



Example: diagnosing stroke patients

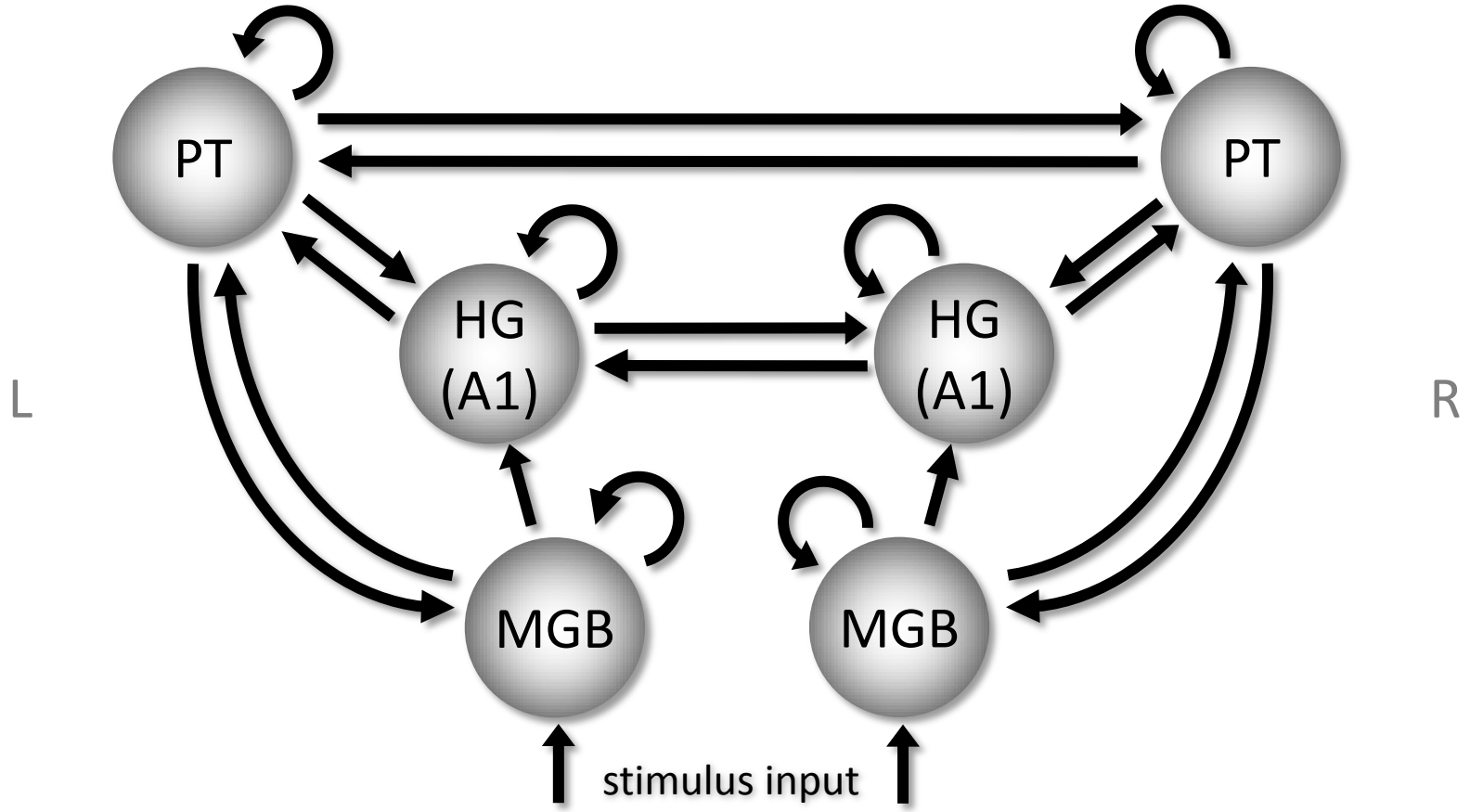


Example: diagnosing stroke patients

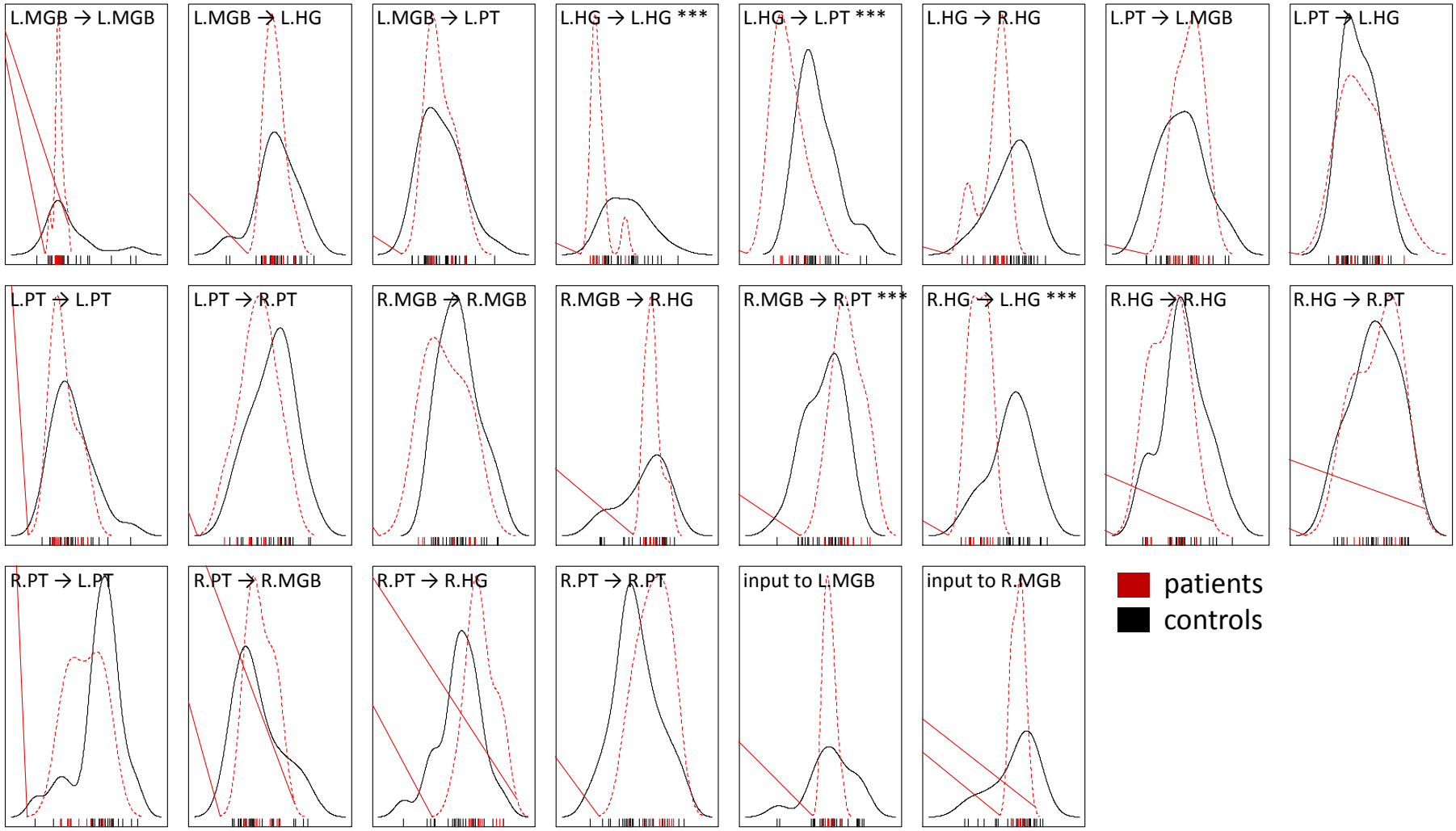


anatomical
regions of interest

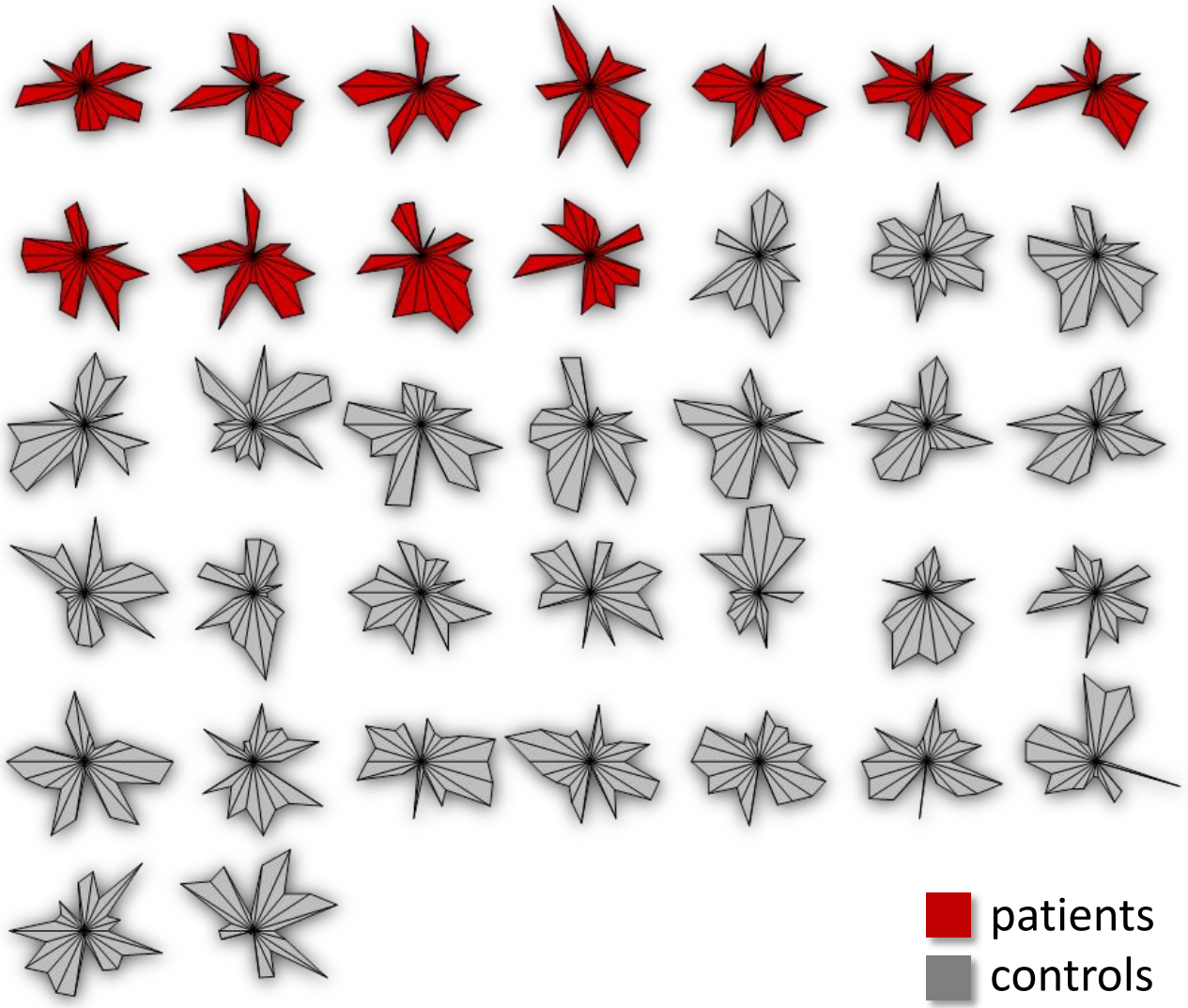
Example: diagnosing stroke patients



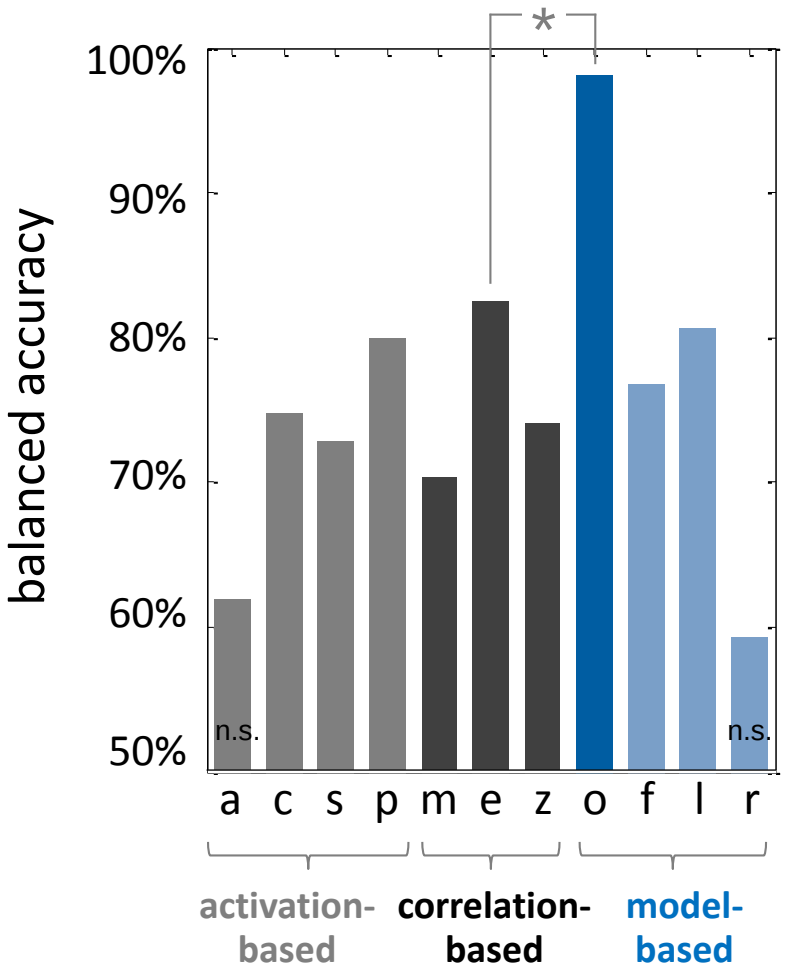
Univariate analysis: parameter densities



Multivariate analysis: connectional fingerprints



Classification performance



Activation-based analyses

- **a** anatomical feature selection
- **c** mass-univariate contrast feature selection
- **s** locally univariate searchlight feature selection
- **p** PCA-based dimensionality reduction

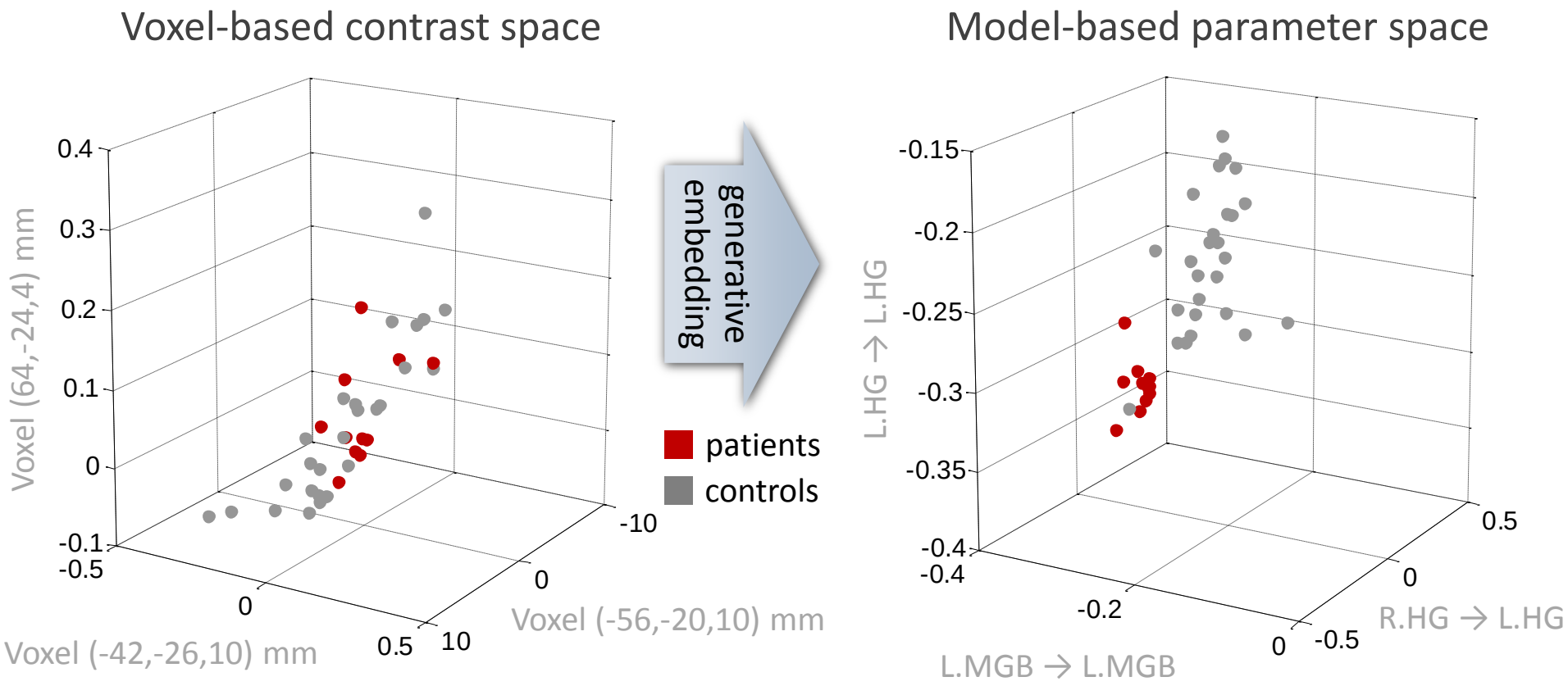
Correlation-based analyses

- **m** correlations of regional means
- **e** correlations of regional eigenvariates
- **z** Fisher-transformed eigenvariates correlations

Model-based analyses

- **o** gen.embed., original full model
- **f** gen.embed., less plausible feedforward model
- **l** gen.embed., left hemisphere only
- **r** gen.embed., right hemisphere only

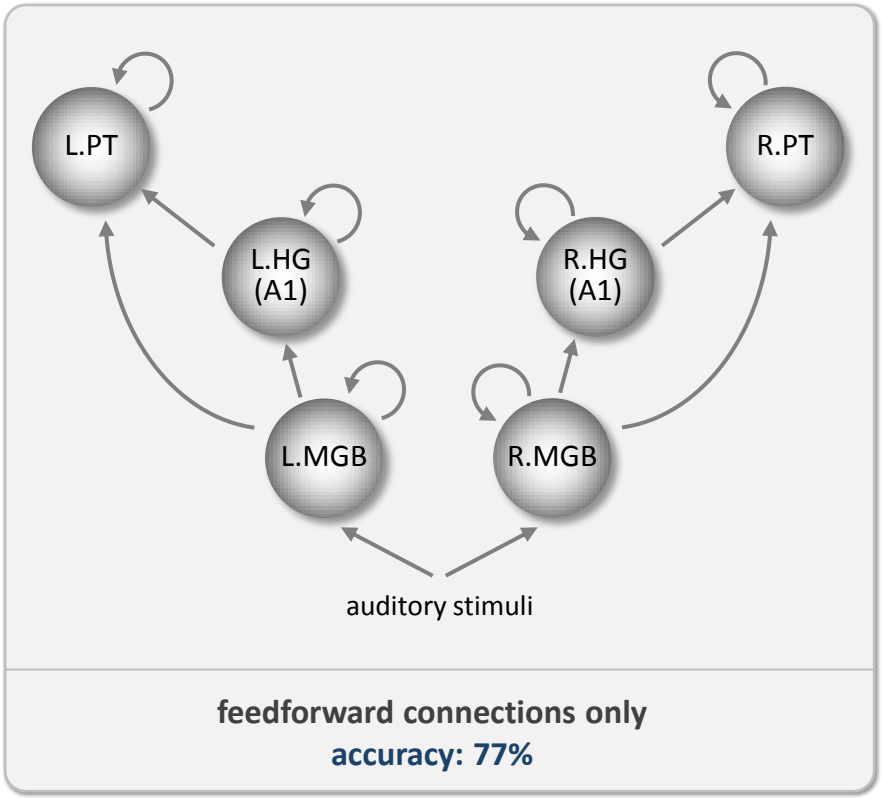
The generative projection



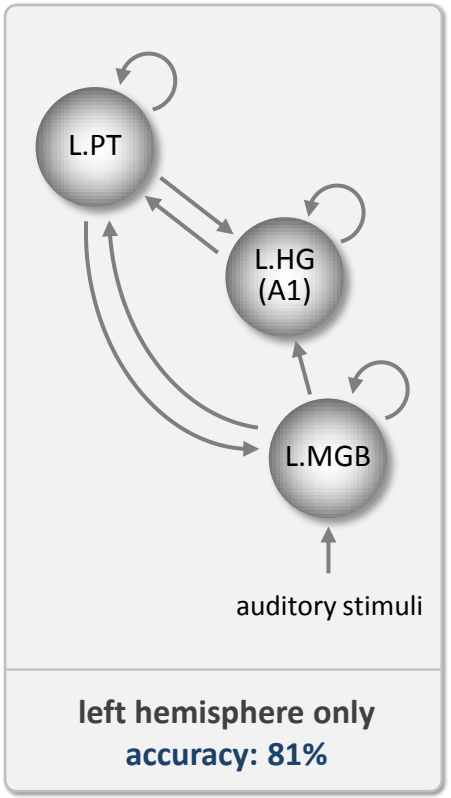
classification accuracy
(using all voxels in the regions of interest)
75%

classification accuracy
(using all 23 model parameters)
98%

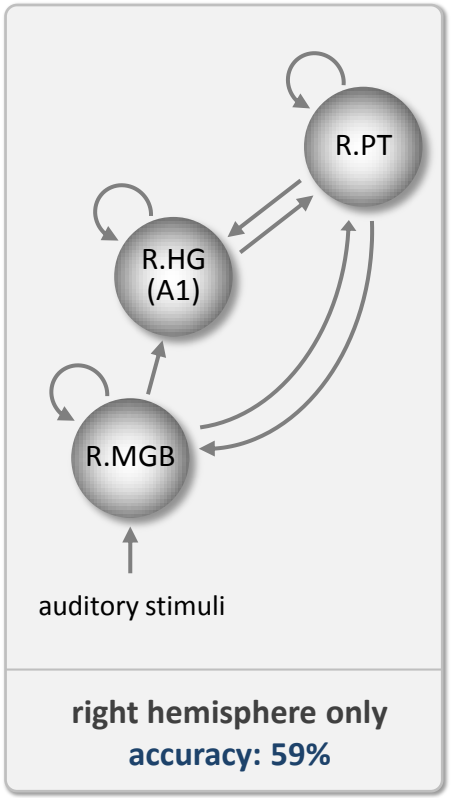
Biologically less plausible models perform poorly



feedforward connections only
accuracy: 77%

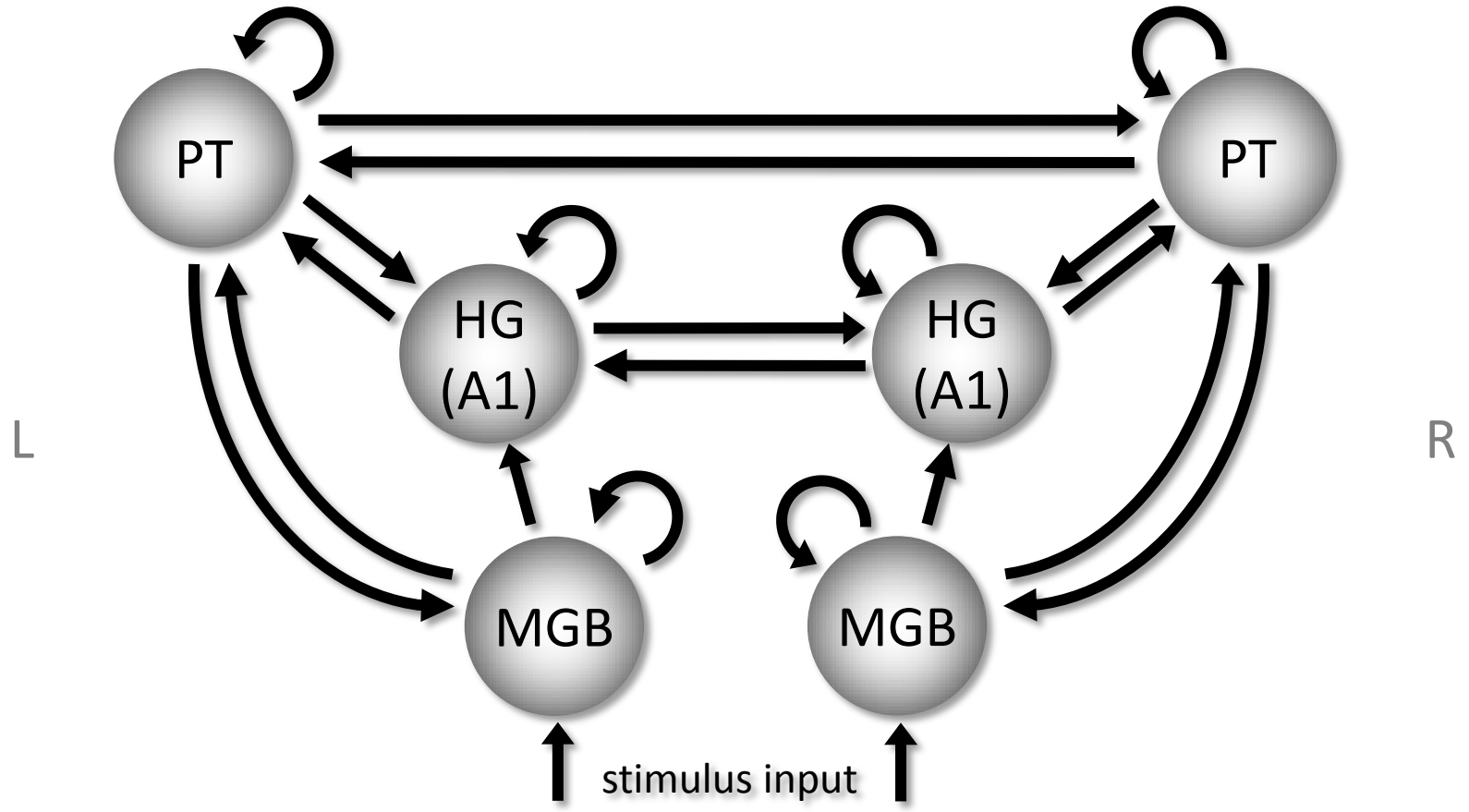


left hemisphere only
accuracy: 81%

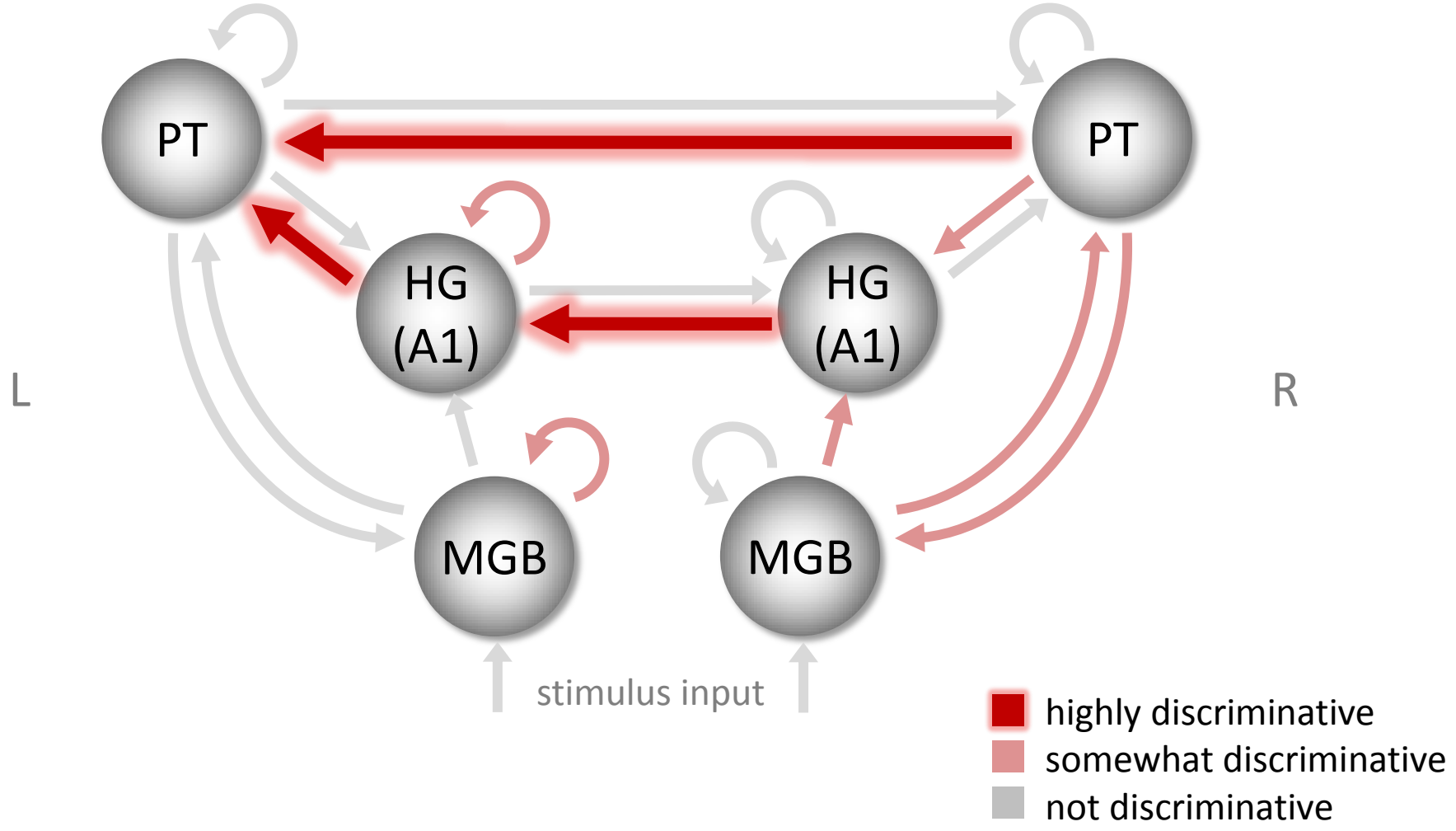


right hemisphere only
accuracy: 59%

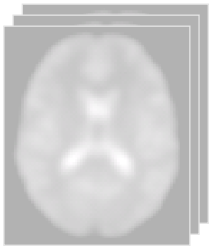
Discriminative features in model space



Discriminative features in model space



Generative embedding and **clustering**

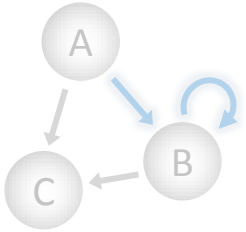


measurements from an individual subject

step 1 — extraction

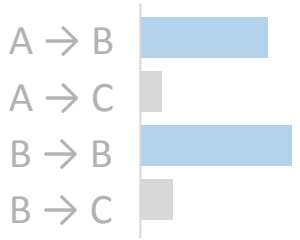
time series in regions of interest

step 2 — modelling



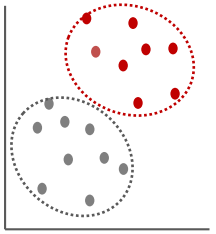
subject-specific generative model

step 3 — embedding



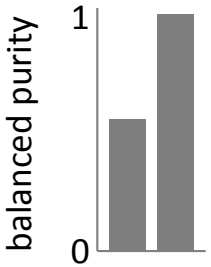
representation in model-based feature space

step 4 — clustering



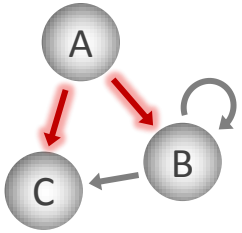
emerging groups of similar subjects?

step 5 — validation



agreement with clinical observations?

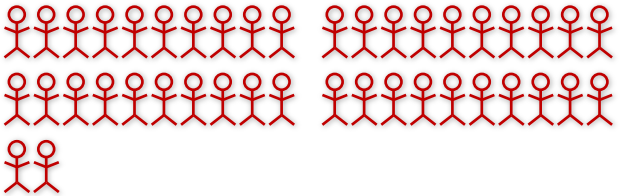
step 6 — interpretation



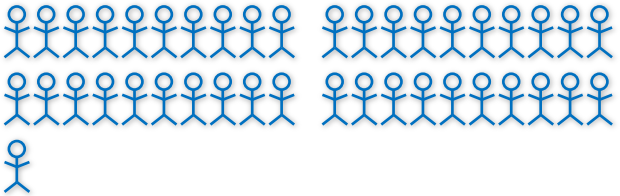
jointly discriminative connection strengths?

Dissecting schizophrenia into subtypes

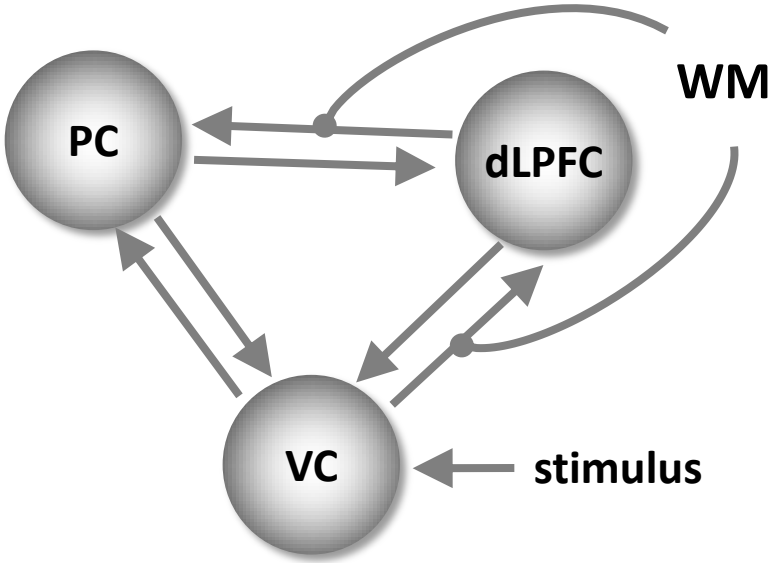
42 patients diagnosed with schizophrenia



41 healthy controls



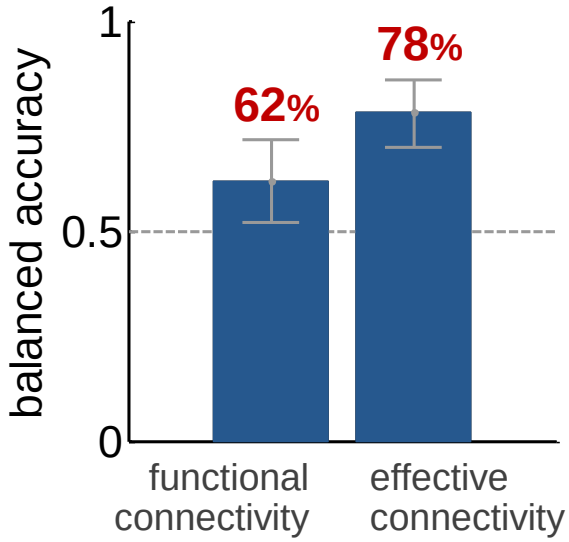
fMRI data acquired during working-memory task & modelled using a three-region DCM



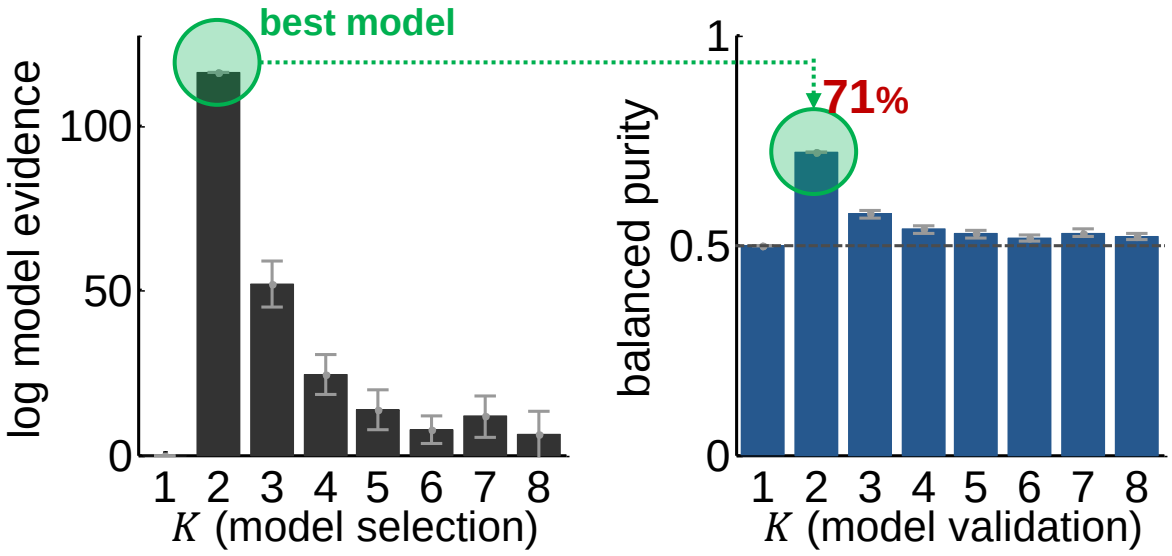
Deserno, Sterzer, Wüstenberg, Heinz, & Schlagenhauf (2012) *J Neurosci*

Distinguishing between schizophrenia and healthy controls

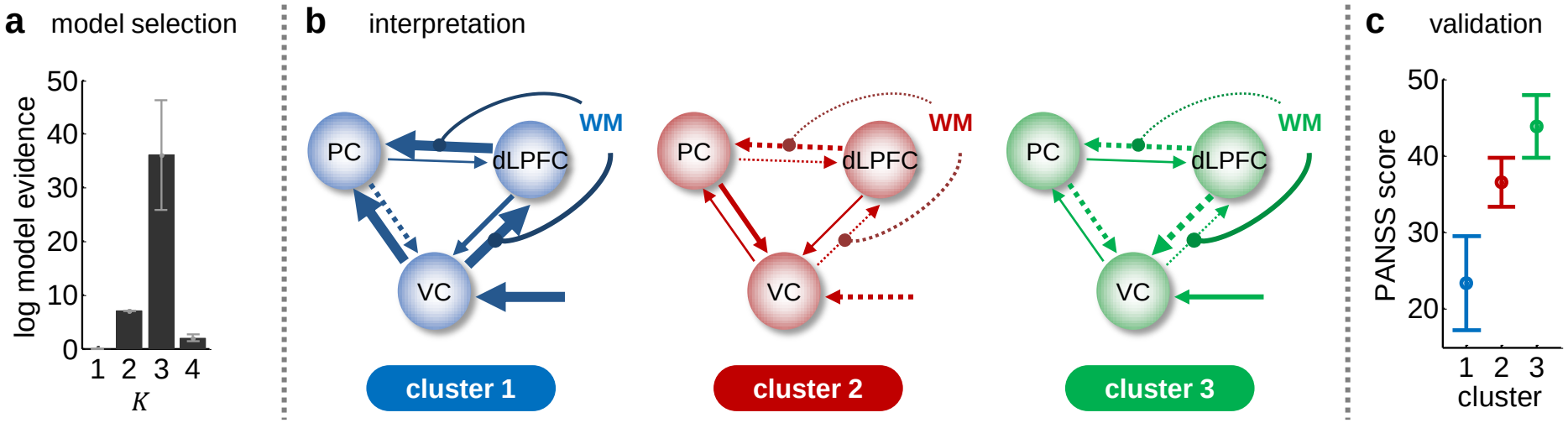
supervised learning:
SVM classification



unsupervised learning:
GMM clustering



Discovering new clinical subtypes

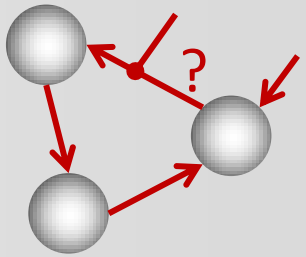


Generative embedding and DCM

Question 1 – What do the data tell us about hidden processes in the brain?

⇨ compute the posterior

$$p(\theta|y,m) = \frac{p(y|\theta,m)p(\theta|m)}{p(y|m)}$$

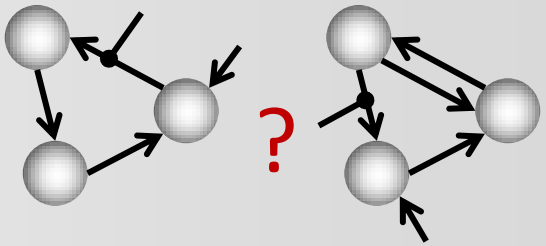


Question 2 – Which model is best w.r.t. the observed fMRI data?

⇨ compute the model evidence

$$p(m|y) \propto p(y|m)p(m)$$

$$= \int p(y|\theta,m)p(\theta|m)d\theta$$

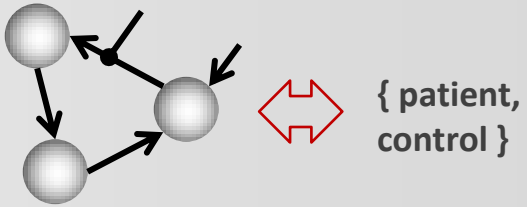


Question 3 – Which model is best w.r.t. an external criterion?

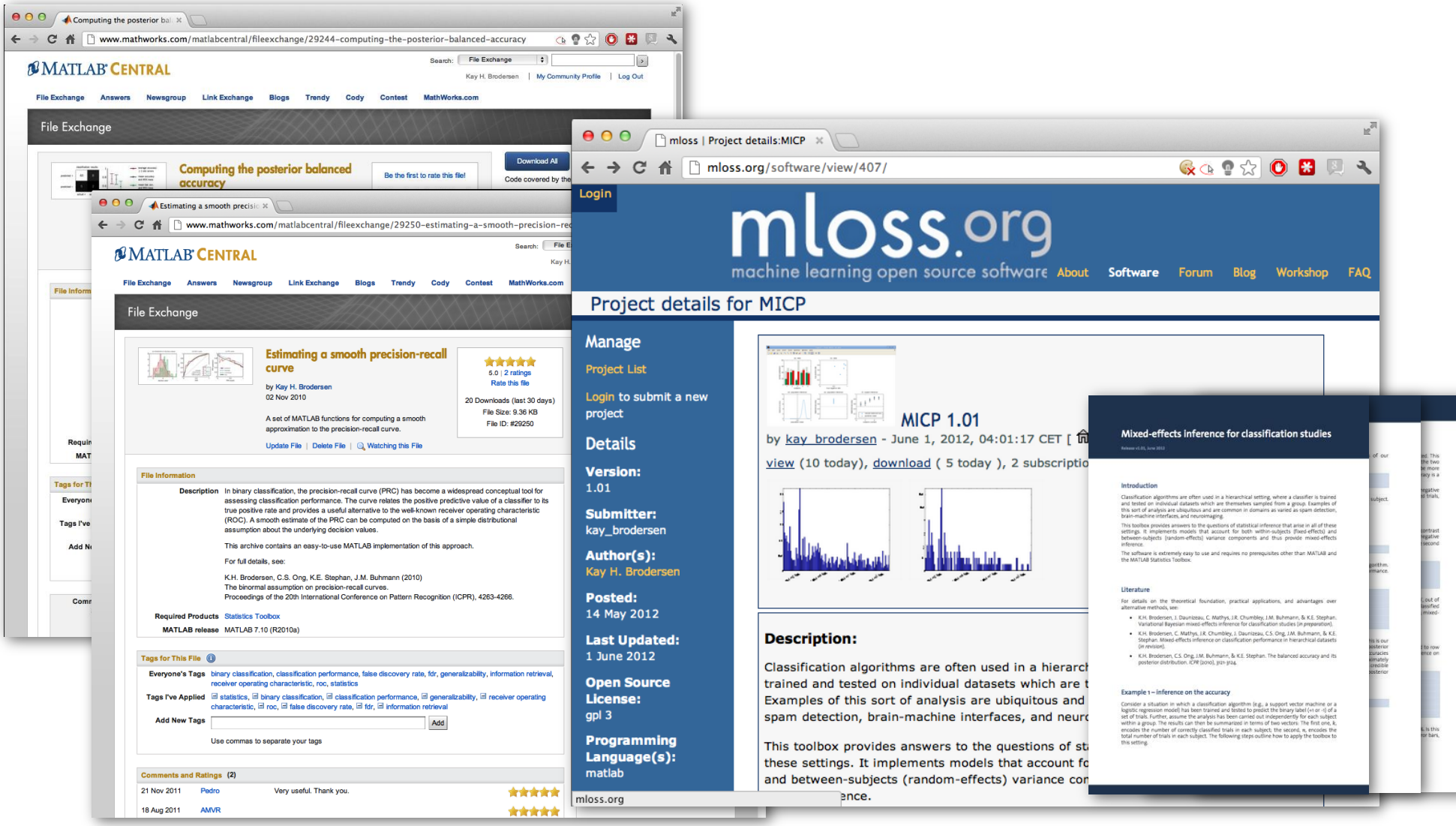
⇨ compute classification accuracy or clustering purity

$$p(h(y) = x|y)$$

$$= \iiint p(h(y) = x|y, y_{\text{train}}, x_{\text{train}})\, p(y)\, p(y_{\text{train}})\, p(x_{\text{train}})\; dy\, dy_{\text{train}}\, dx_{\text{train}}$$



Toolbox releases



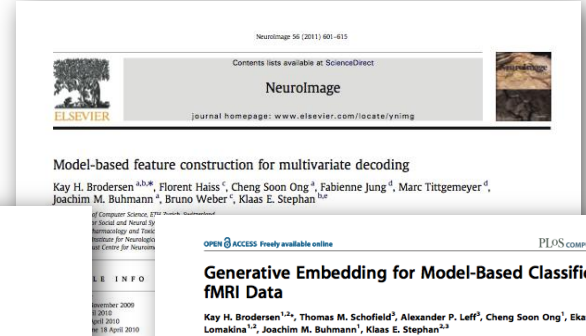
Accepted first-author manuscripts

2010 International Conference on Pattern Recognition

The binormal assumption on precision-recall curves

Kay H. Brodersen*†, Cheng Soon Ong*, Klaas E. Stephan† and Joachim M. Buhmann*

*Department of Computer Science, ETH Zurich, Switzerland; kay.brodersen@inf.ethz.ch
†Institute for Empirical Research in Economics, University of Zurich, Switzerland



NeuroImage 56 (2011) 601–615

Contents lists available at ScienceDirect

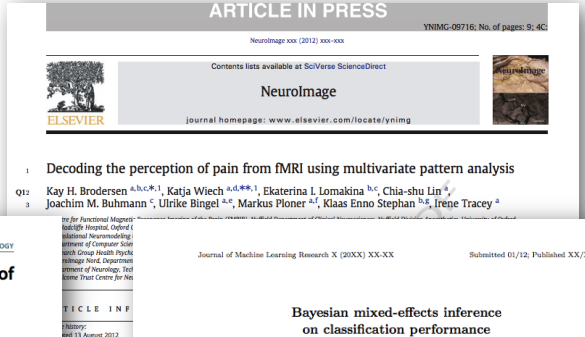
NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

ELSEVIER

Model-based feature construction for multivariate decoding

Kay H. Brodersen [a,b,*], Florent Haiss [c], Cheng Soon Ong [a], Fabienne Jung [d], Marc Tittgemeyer [d], Joachim M. Buhmann [a], Bruno Weber [c], Klaas E. Stephan [b,e]



ARTICLE IN PRESS

YNIMG-09716; No. of pages: 9; 4C

NeuroImage xxx (2012) xxx–xxx

Contents lists available at SciVerse ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

ELSEVIER

Decoding the perception of pain from fMRI using multivariate pattern analysis

Kay H. Brodersen [a,b,c,*,1], Katja Wiech [a,d,**,1], Ekaterina I. Lomakina [b,c], Chia-shu Lin [a], Joachim M. Buhmann [c], Ulrike Bingel [b,e], Markus Ploner [b,f], Klaas Enno Stephan [b,g], Irene Tracey [a]

2010 International Conference on Pattern Recognition

The balanced accuracy and its posterior distribution

Kay H. Brodersen*†, Cheng Soon Ong*, Klaas E. Stephan† and Joachim M. Buhmann*

*Department of Computer Science, ETH Zurich, Switzerland; kay.brodersen@inf.ethz.ch
†Institute for Empirical Research in Economics, University of Zurich, Switzerland

OPEN ACCESS Freely available online

PLoS COMPUTATIONAL BIOLOGY

Generative Embedding for Model-Based Classification of fMRI Data

Kay H. Brodersen [1,2*], Thomas M. Schofield [3], Alexander P. Leff [3], Cheng Soon Ong [1], Ekaterina I. Lomakina [1,2], Joachim M. Buhmann [1], Klaas E. Stephan [2,3]

1 Department of Computer Science, ETH Zurich, Zurich, Switzerland, 2 Laboratory for Social and Neural Systems Research, Department of Economics, University of Zurich, Zurich, Switzerland, 3 Wellcome Trust Centre for Neuroimaging, University College London, London, United Kingdom

Journal of Machine Learning Research X (20XX) XX-XX

Submitted 01/12; Published XX/XX

Bayesian mixed-effects inference on classification performance in hierarchical datasets

Kay H. Brodersen [1,2,3] BRODERSEN@BIOMED.EE.ETHZ.CH
Christoph Mathys [1,3] CHMATHYS@ETHZ.CH
Justin R. Chumbley [3] JUSTIN.CHUMBLEY@ECON.UZH.CH
Jean Daunizeau [1,3] J.DAUNIZEAU@FIL.ION.UCL.AC.UK
Cheng Soon Ong [2] CHENGSOON.ONG@UNIMELB.EDU.AU
Joachim M. Buhmann [2] JBUHMANN@INF.ETHZ.CH
Klaas E. Stephan [1,3,4] STEPHAN@BIOMED.EE.ETHZ.CH

[1] Translational Neuromodeling Unit (TNU), Institute for Biomedical Engineering, University of Zurich & ETH Zurich, Switzerland
[2] Machine Learning Laboratory, Department of Computer Science, ETH Zurich, Switzerland
[3] Laboratory for Social and Neural Systems Research (SNS), Department of Economics, University of Zurich, Switzerland
[4] Wellcome Trust Centre for Neuroimaging, University College London, United Kingdom

Editor: Aapo Hyvarinen

Abstract

Classification algorithms are frequently used on data with a natural hierarchical structure. For instance, classifiers are often trained and tested on trial-wise measurements, separately for each subject within a group. One important question is how classification outcomes observed in individual subjects can be generalized to the population from which the group was sampled. To address this question, this paper introduces novel statistical models that are guided by three desiderata. First, all models explicitly respect the hierarchical nature of the data, that is, they are mixed-effects models that simultaneously account for within-subjects (fixed-effects) and across-subjects (random-effects) variance components. Second, maximum-likelihood estimation is replaced by full Bayesian inference in order to enable natural regularization of the estimation problem and to afford conclusions in terms of posterior probability statements. Third, inference on classification accuracy is complemented by inference on the balanced accuracy, which avoids inflated accuracy estimates for imbalanced datasets. We introduce hierarchical models that satisfy these criteria and demonstrate their advantages over conventional methods using MCMC implementations for model inversion and model selection on both synthetic and empirical data. We envisage that our approach will improve the sensitivity and validity of statistical inference in future hierarchical classification studies.

Keywords: beta-binomial, normal-binomial, balanced accuracy, Bayesian inference, group studies

Reception

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

ETH Life

ETH Life » News' Archive » Brain Network Reveals Disorders | **Deutsch** | Print | Send this page |

Home
Topics
All articles
Science
Campus
Columns
Globetrotter
Surprise
ETH in-house
Dossiers
Events
ETH Life Archive
Contact
Search results

DOSSIERS

Health Sciences and Technologies

Nanotechnology

EU flagship programme

NEWSLETTER

Subscribe
Unsubscribe / Change

READER COMMENTS

11.09.12 QS-Ranking: Zweifelhafter Wert von Rankings

06.09.12 Kolumne Rahel Zoller: Abfall - Als Zeichen der Akzeptanz ?

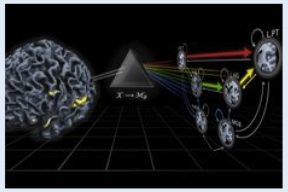
29.08.12 Atemtraining und

Published: 12.04.12

Science

Brain Network Reveals Disorders

Researchers at ETH Zurich and the University of Zurich identify a new method of unerringly detecting the presence of pathophysiological changes in the brain. The new method was developed in order to gain a mechanistic understanding of schizophrenia and other spectrum disorders, which will lead to more accurate diagnoses and more effective treatments.



Brain model (left) depicting brain activity stimulated by speech processing (yellow). The new method allows for the mathematical modeling of interactions between regions within the brain (right). The prism represents the transition or "Generative Embedding." (Image: Brodersen KH/ ETH Zurich) (gallery)

When mathematical genius John Nash was diagnosed with schizophrenia, the chance for recovery was slim. Medicine in the 1960's simply had no convincing explanations for his condition. Alarmingly, things don't look much better nowadays: depression, addiction, schizophrenia, and other spectrum disorders remain among the toughest challenges for medicine. This is because they are caused by complicated and largely unknown interactions between genes and the environment. Different disease mechanisms may underlie similar, or even identical, symptoms. This means that the effect of any given drug may vary hugely across individuals, resulting in trial-and-error treatment. In addition, conditions whose biological basis is not well-understood may be perceived as particularly stigmatizing.

Most spectrum disorders lack a physiological definition altogether; they are simply described in terms of particular symptoms. This is problematic when these symptoms are caused by different disease mechanisms. Conversely, existing disease classifications frequently group patients with disjoint symptoms under the same label: a person with delusions and disorganized thought, for instance, can be diagnosed with schizophrenia, just as somebody else suffering from hallucinations and movement problems. Examples such as this one show that the development of more specific diagnoses and more effective treatment will require a mechanistic understanding of the pathophysiological mechanisms underlying spectrum disorders.

One step in this direction has recently been made by Kay Henning Brodersen and Klaas Enno Stephan at ETH Zurich and the University of Zurich. Within the

SCIENCE

more article

CAMPUS

more article

GLOBETR

more article

ETH IN-H

more article
house

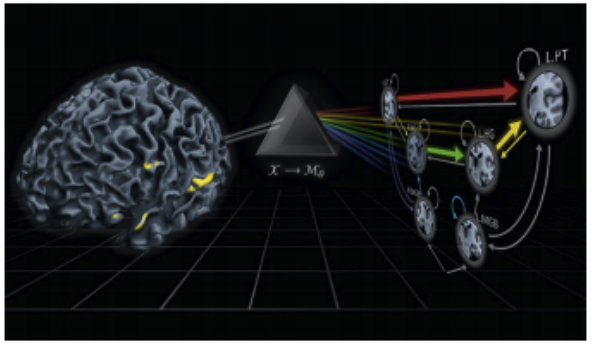
SystemsX.ch Newsletter #24
June 2012

X-Letter 24

Translational neuromodeling: from imaging science to clinical applications

When mathematical genius John Nash was diagnosed with schizophrenia, the chance for a quick recovery was slim. Medicine in the 1960's simply had no convincing explanations for his condition. Alarmingly, things don't look much better nowadays: depression, addiction, schizophrenia, and other spectrum disorders remain among the toughest challenges for medicine. This is because they are caused by complicated and largely unknown interactions between genes and the environment. Different disease mechanisms may underlie similar, or even identical, symptoms. This means that the effect of any given drug may vary hugely across individuals, resulting in trial-and-error treatment. In addition, conditions whose biological basis is not well-understood may be perceived as particularly stigmatizing.

They analysed brain activity from two groups of participants: one group of stroke patients that suffered from language impairments; and one group of healthy volunteers. While undergoing functional magnetic resonance imaging



Mathematical microscopes could help dignose psychiatric disorders.

show that the development of more specific diagnoses and more effective treatment will require a better understanding of the pathophysiological mechanisms underlying spectrum disorders.

One step into this direction has recently been taken by Kay Henning Brodersen and Klaas Enno Stephan of ETH Zurich and the University of Zurich. Within the framework of the SystemsX.ch project 'Neurochoice', the two researchers investigate how insights

(fMRI), participants were asked to passively listen to speech. A mathematical model was then used to assess, separately within each participant, how brain regions involved in speech processing interacted. Notably, none of the brain regions included in the model had been affected by the stroke in the patients. The researchers then asked whether it was possible to automatically detect the presence of a remote lesion from patterns of brain connectivity in the healthy part of the brain. "Using our

More specific diagnosis and effective treatment

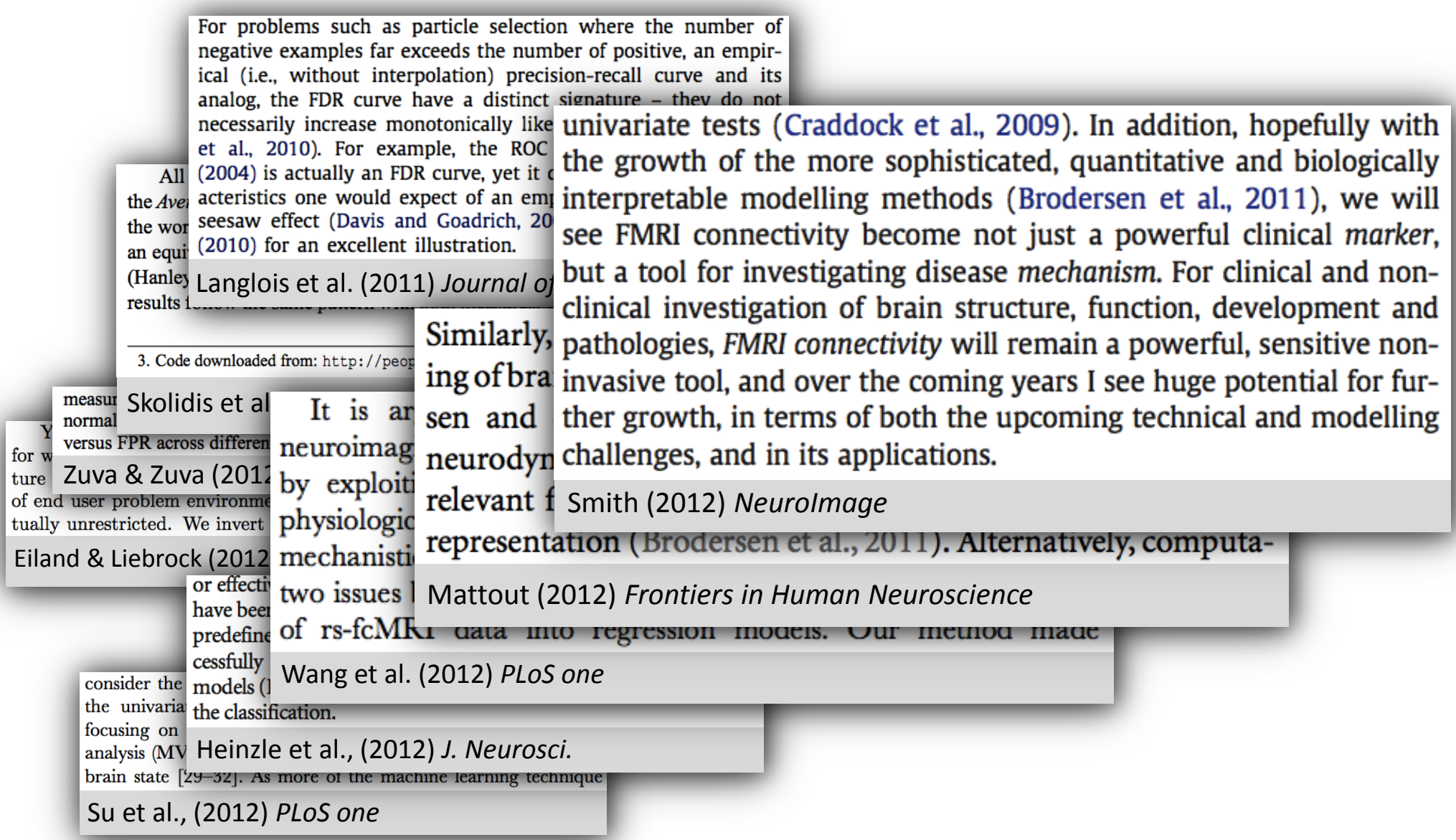
Most spectrum disorders lack a physiological definition altogether; they are simply described in terms of particular symptoms. This is problematic when these symptoms are caused by different disease mechanisms. Conversely, existing disease classifications frequently group patients with disjoint symptoms model of brain function, we were able to diagnose patients with an accuracy of 98%," says Brodersen, first author of the study. "This became possible by tying together dynamic causal models of neuronal dynamics with mathematical techniques from machine learning and Bayesian inference."

Further investigations in other patients planned

In contrast to subtle spectrum disorders, of course, this initial proof-of-principle study concerned a rather salient clinical condition, that is, language impairments caused by a stroke. In the future, Stephan and Brodersen therefore plan to investigate whether their approach might work equally well for those diseases where contemporary medicine is struggling, such as schizophrenia, depression, and addiction. The two researchers hope that their approach will help dissect these spectrum disorders into pathophysiologically well-defined subgroups. Identifying such subgroups would provide an important step towards more specific diagnoses and may eventually predict the most effective treatment for an individual patient.

K. H. Brodersen, T. M. Schofield, A. P. Leff, C. S. Ong, E. I. Lomakina, J. M. Buhmann, K. E. Stephan (2011). Generative embedding for model-based classification of fMRI data. PLoS Computational Biology, 7(6): e1002079.

Reception


For problems such as particle selection where the number of negative examples far exceeds the number of positive, an empirical (i.e., without interpolation) precision-recall curve and its analog, the FDR curve have a distinct signature – they do not necessarily increase monotonically like
et al., 2010). For example, the ROC
(2004) is actually an FDR curve, yet it
acteristics one would expect of an emp
seesaw effect (Davis and Goadrich, 20
(2010) for an excellent illustration.
Langlois et al. (2011) Journal of
univariate tests (Craddock et al., 2009). In addition, hopefully with the growth of the more sophisticated, quantitative and biologically interpretable modelling methods (Brodersen et al., 2011), we will see FMRI connectivity become not just a powerful clinical marker, but a tool for investigating disease mechanism. For clinical and non-clinical investigation of brain structure, function, development and pathologies, FMRI connectivity will remain a powerful, sensitive non-invasive tool, and over the coming years I see huge potential for further growth, in terms of both the upcoming technical and modelling challenges, and in its applications.
Smith (2012) NeuroImage
Skolidis et al
Zuva & Zuva (2012
Eiland & Liebrock (2012
Similarly,
ing of bra
sen and
neurodyn
relevant f
representation (Brodersen et al., 2011). Alternatively, computa-
Mattout (2012) Frontiers in Human Neuroscience
It is ar
neuroimag
by exploiti
physiologic
mechanisti
two issues
of rs-fcMRI data into regression models. Our method made
Wang et al. (2012) PLoS one
the classification.
Heinzle et al., (2012) J. Neurosci.
Su et al., (2012) PLoS one

Thank you

Supervisors

Joachim M. Buhmann

Klaas E. Stephan

Colleagues and collaborators

Paul Allen

Alberto Busetto

Morteza Chehreghani

Justin Chumbley

Jean Daunizeau

Lorenz Deserno

Andreea Diaconescu

Ajita Gupta

Florent Haiss

Jakob Heinzle

Sandra Iglesias

Fabienne Jung

Lars Kasper

Laura Hughes

Alex Leff

Falk Lieder

Chia-shu Lin

Zhihao Lin

Kate Lomakina

Christoph Mathys

Philip McGuire

Rosalyn Moran

Cheng Soon Ong

Saee Paliwal

Will Penny

Ulrich Pfeiffer

James Rowe

Vera Schäfer

Leo Schillbach

Florian Schlagenhauf

Tom Schofield

Gabor Stefanics

Marc Tittgemeyer

Irene Tracey

Kai Vogeley

Bruno Weber

Katja Wiech

Administration

Eva Gschwend

Rita Klute

Ursi Meier

Denise Spicher

Suzanne Wilde

Karin Zeder