Mobile real-time EEG imaging

Lars Kai Hansen
DTU Informatics
Technical University of Denmark
lkh@imm.dtu.dk

Co-workers:

Arek Stopczynski, Carsten Stahlhut, Jakob E. Larsen, Michael K. Petersen, Sofie T. Hansen, Ivana Konvalinka






Cimbi
Center for integrated molecular brain imaging

DTU

OUTLINE

Mobile real-time EEG Imaging

Why real-time imaging?
Current implementation

Sparse mean field method ”Variational Garrote”
vs Lasso/ARD
Temporal smoothness Markov prior

Where do we want to go with
mobile solutions?

Smartphone Brain Scanner

Fig. 1. Handheld brain scanner components. Emotiv EPOC wireless EEG headset (1), Emotiv Receiver module with USB connector (2), USB connector and adapter (3+4), and Nokia N900 mobile phone. The total cost of the system is less than USD1000.

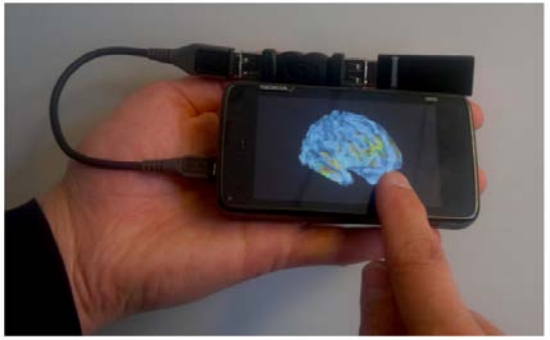


Fig. 3. A user interacting with a 3D model of the brain using the handheld brain scanner device with touch-based interaction.

A. Stopczynski, J.E. Larsen, C. Stahlhut, M.K. Petersen, L.K. Hansen.
A Smartphone Interface for a Wireless EEG Headset with Real-Time 3D Reconstruction.
In Proc. Affective Computing and Intelligent Interaction. Springer Lecture Notes in Computer Science 6975: 317-318 (2011).

Mobilizing personal state decoding

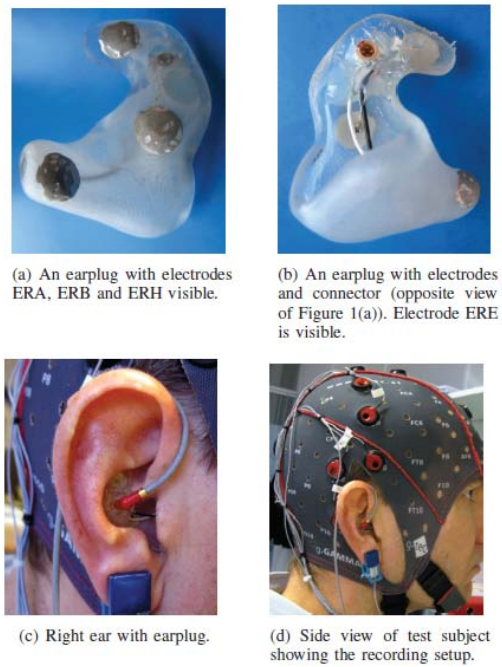
More naturalistic conditions for brain imaging experiments

Long time observations in the wild: 24/7 monitoring - ”quantified self”

EEG real time 3D imaging for bio-feedback

Illustration of HypoSafe implantable device

hypo safe®



(a) An earplug with electrodes ERA, ERB and ERH visible.

(b) An earplug with electrodes and connector (opposite view of Figure 1(a)). Electrode ERE is visible.

(c) Right ear with earplug.

(d) Side view of test subject showing the recording setup.

Fig. 1. View of a right ear earplug and the Ear-EEG recording setup.

P. Kidmose et al.

Auditory Evoked Responses from Ear-EEG Recordings.

IEEE EMBS (2012)

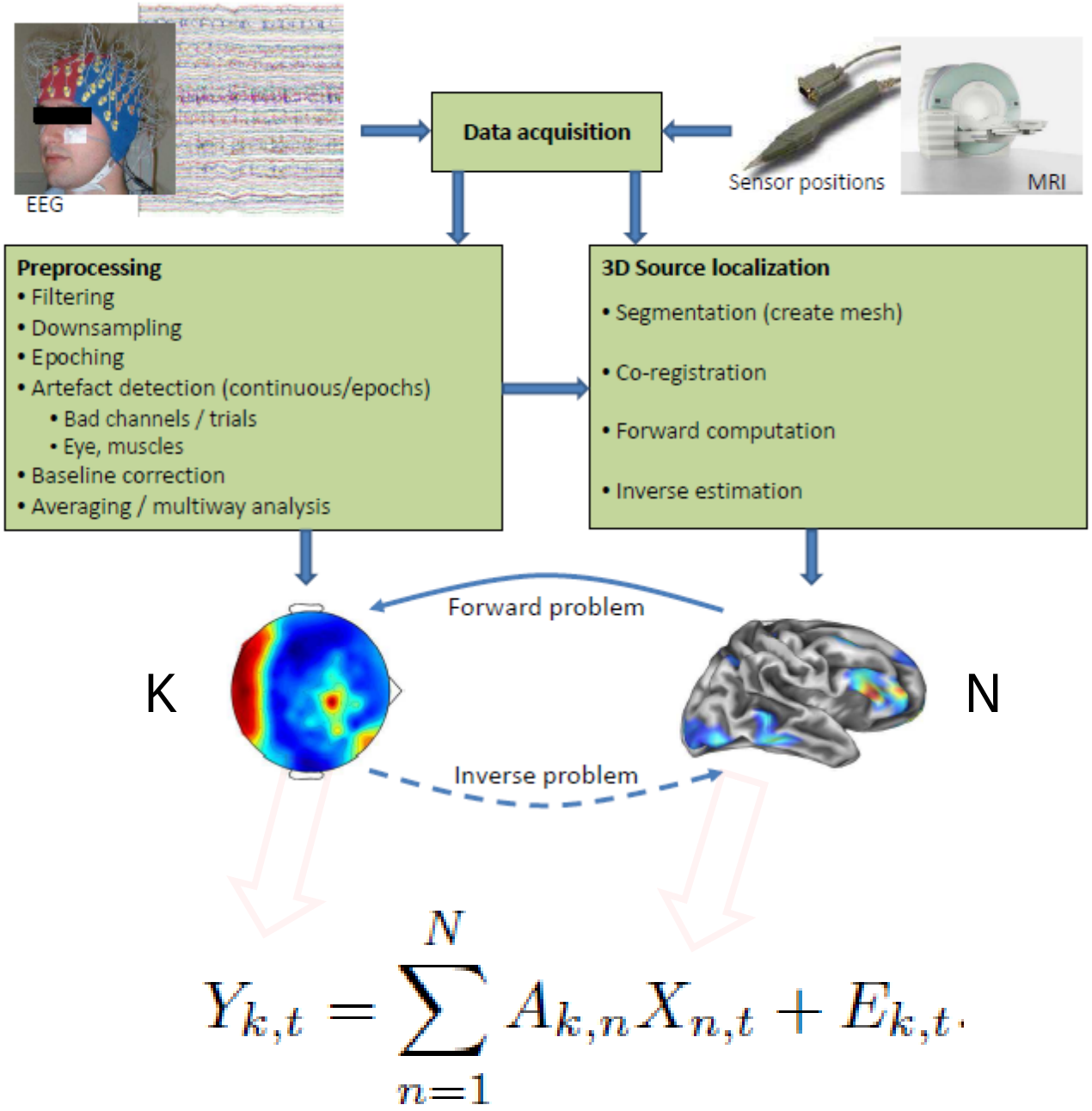
EEG imaging

Linear ill-posed inverse problem

X: N x T
Y: K x T
A: K x N

N >> K

Need priors to solve!



$$Y_{k,t} = \sum_{n=1}^{N} A_{k,n} X_{n,t} + E_{k,t}.$$

C. Stahlhut: Functional Brain Imaging by EEG: A Window to the Human Mind. PhD-Thesis (2011), DTU Informatics

Intermezzo

Multimodal integration by letting one modality act as prior for the other

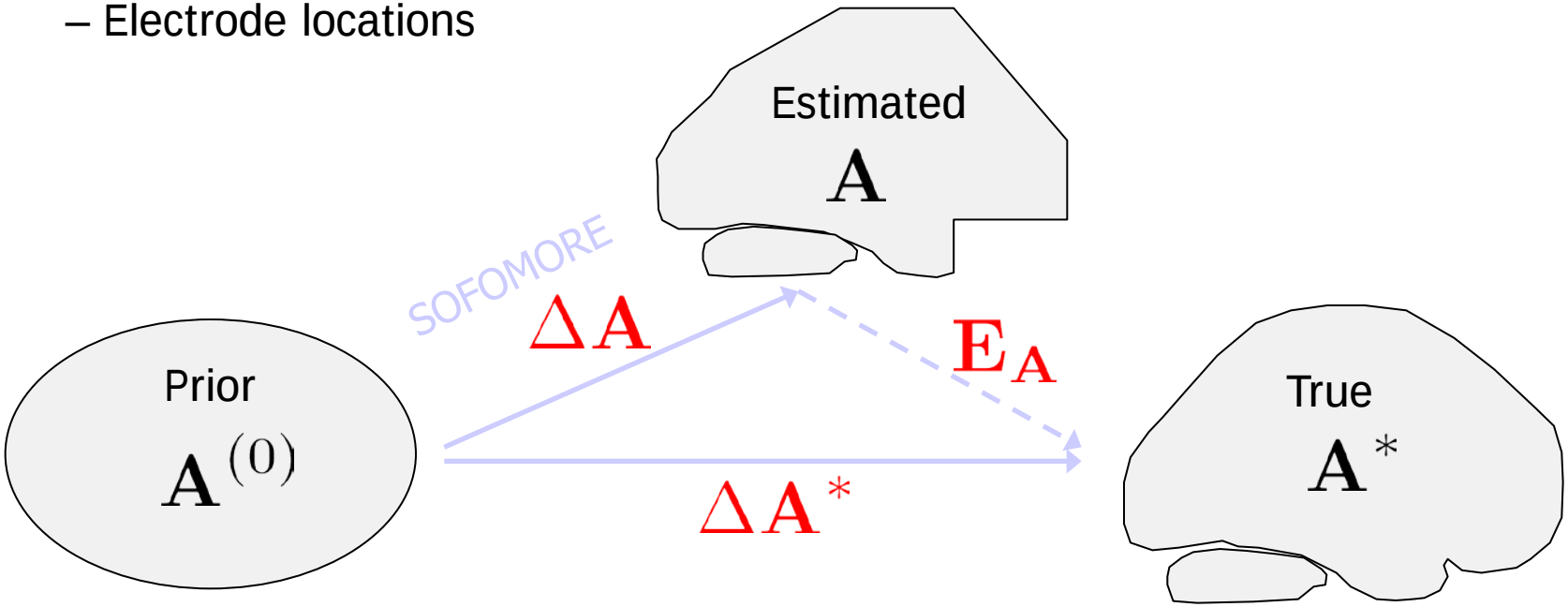
Here:

The EEG forward model is inaccurate...but useful as ”prior”

Reconstruction of the forward model

Uncertainties involved in the estimation of the forward model

- Tissue segmentation
- Tissue conductivities
- Electrode locations



Previous work:

- (Lew et al.,2007; Plis et al., 2007)

Experiments: EEG face-evoked response

SOFOMORE Model

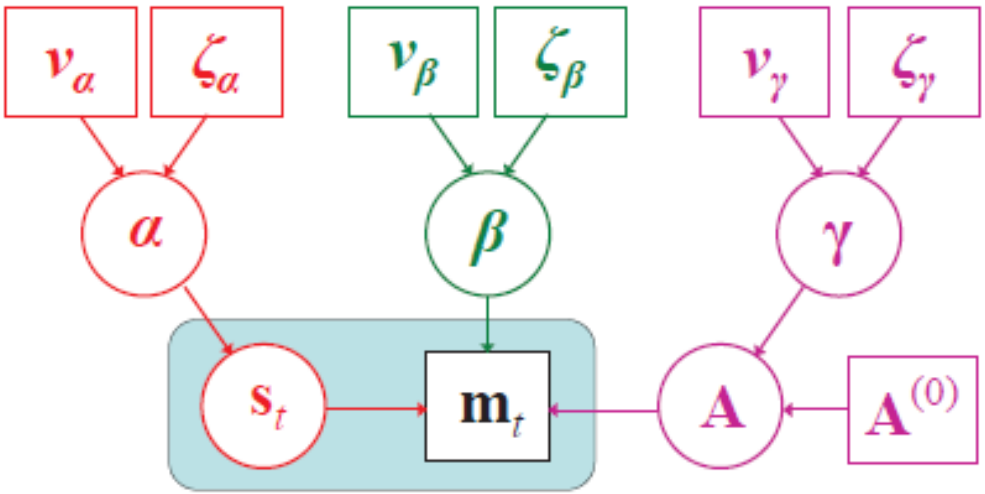


Figure 4.2: **Graphical representation of the SOFOMORE model.** The blue box including the sources $\mathbf{s}_t$ and observations $\mathbf{m}_t$ indicates expansion over time t. At the lowest level in the hierarchical structure we also find the forward model $\mathbf{A}$ with fixed prior mean $\mathbf{A}^{(0)}$. The middle layer includes $\boldsymbol{\alpha}$ precision parameter for the sources with a seperate precision parameter (inverse variance) assigned to each dipole. β is the inverse variance of the noise contribution and $\boldsymbol{\gamma}$ includes a precision parameter to each column in $\mathbf{A}$. At the top level we have the hyperhyperparameters controlling the hyperparameters in the middle layer.

Cimbi Center for integrated molecular brain imaging

C. Stahlhut, M. Mørup, O. Winther, L.K. Hansen.
Simultaneous EEG Source and Forward Model Reconstruction (SOFOMORE) using a Hierarchical Bayesian Approach.
Journal of Signal Processing Systems, 65(3):431-444 (2011).

Effect of wrong forward model

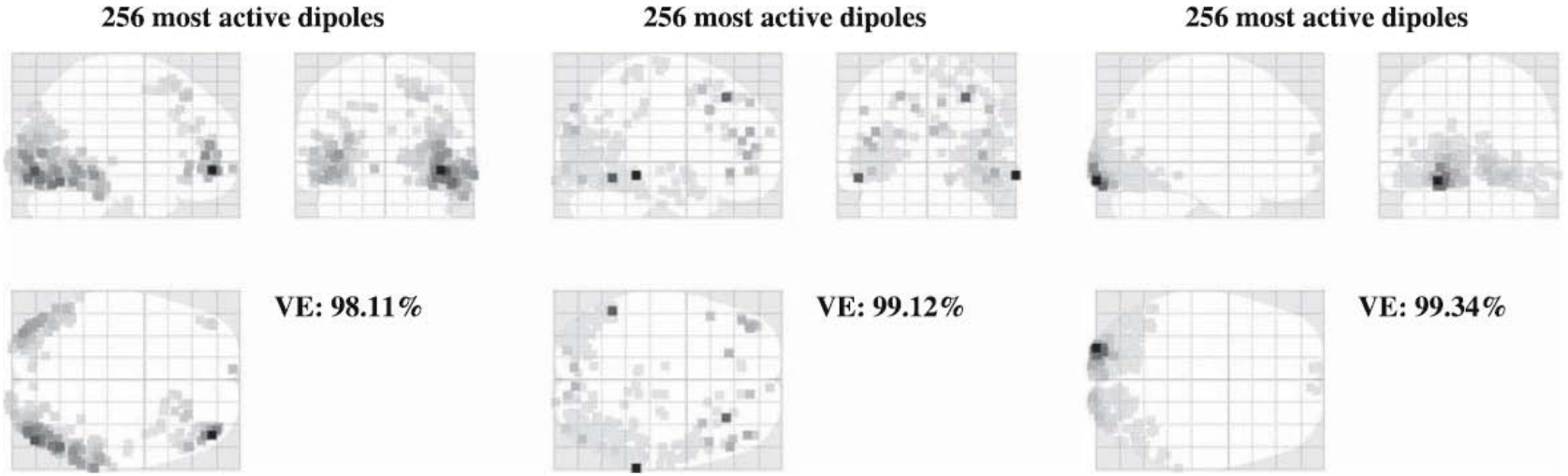


Figure 9 Estimated activity at $t = 170$ ms after stimulus. Tissue conductivities brain:skull:scalp = 0.33:0.0041:0.33 S/m are used. Activity in the left and right occipital region is estimated by MN with the primary activity located in the right occipital region. Moreover, right frontal activity is reconstructed. The ARD leads to quite scattered activity with two dominating dipoles located in the left and right temporal lope. SOFOMORE reconstructs activity both in the left and right visual cortex with dominating activity in the left region.

Why 3D real-time imaging?

Enable on-line visual quality control

Neurofeed applications can be based on activity in specific brain structures /networks

Context priors may relate to 3D location (from meta analysis)

Evidence that BCI /decoding can be improved by 3D representation

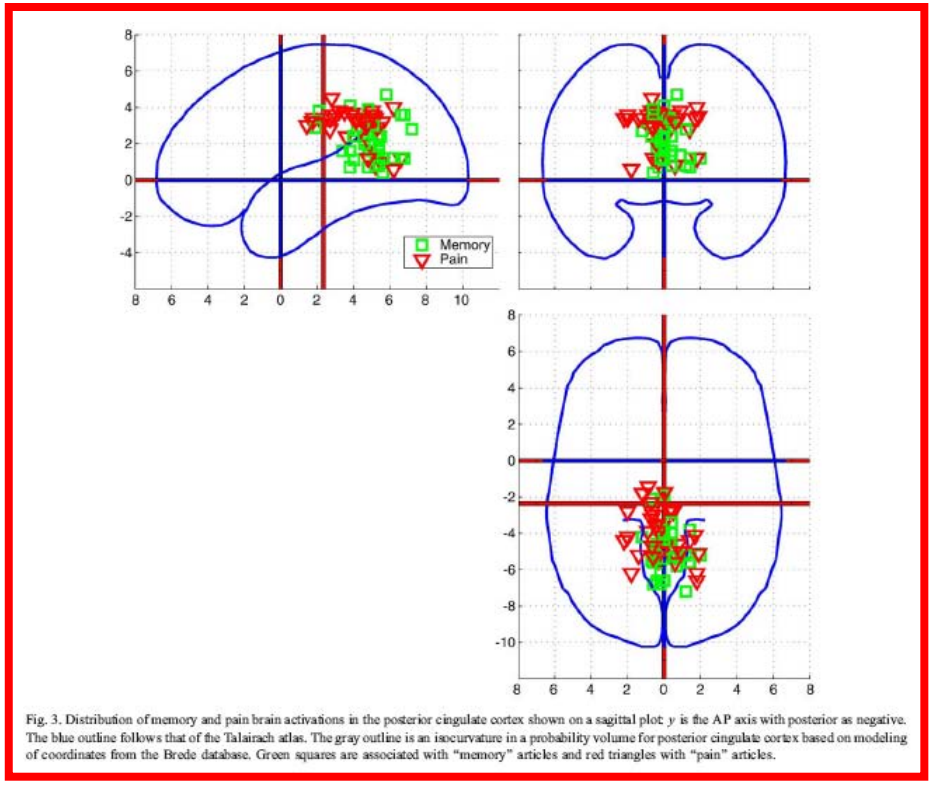


Fig. 3. Distribution of memory and pain brain activations in the posterior cingulate cortex shown on a sagittal plot: y is the AP axis with posterior as negative. The blue outline follows that of the Talairach atlas. The gray outline is an isocurvature in a probability volume for posterior cingulate cortex based on modeling of coordinates from the Brede database. Green squares are associated with "memory" articles and red triangles with "pain" articles.

Finn Årup Nielsen, Daniela Balslev, Lars Kai Hansen, "Mining the Posterior Cingulate: Segregation between memory and pain components". NeuroImage, 27(3):520-532, (2005)

Trujillo-Barreto, Nelson J., Eduardo Aubert-Vázquez, and Pedro A. Valdés-Sosa. "Bayesian model averaging in EEG/MEG imaging." NeuroImage 21, no. 4 (2004): 1300-1319.

Source representation can improve decoding

Besserve et al. (2011)

… reconstructing the underlying cortical network dynamics significantly outperforms a usual electrode level approach in terms of information transfer and also reduces redundancy between coherence and power features, supporting a decrease of volume conduction effects. Additionally, the classifier coefficients reflect the most informative features of network activity, showing an important contribution of localized motor and sensory brain areas, and of coherence between areas up to 6 cm distance.

Ahn et al. (2012)

… source imaging may enable noise filtering, and in so doing, make some invisible discriminative information in the sensor space visible in the source space.

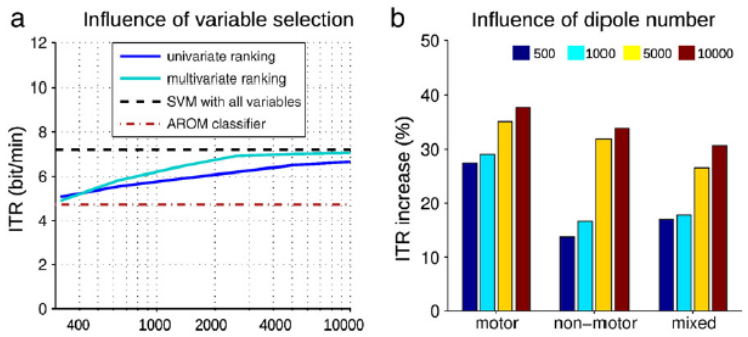


Fig. 6. Effect of reducing the number of sources or variables, for *power+coherence* quantification at the source level. a) Average ITR as a function of the number of variables for two variable ranking techniques: univariate ranking with a Student's *t*-test and multivariate ranking with the coefficient of a SVM classifier. The ITR values using a sparse number of variables with the AROM classifier (see text) and all variables with an SVM are plotted for comparison. b) Influence of the number of cortical dipoles used in the forward model on the ITR: percentage improvement of ITR with respect to electrode level quantification, for each type of couples of tasks (motor, non-motor and mixed couples).

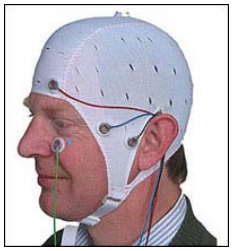
Congedo, Marco, Fabien Lotte, and Anatole Lécuyer. "Classification of movement intention by spatially filtered electromagnetic inverse solutions." *Physics in Medicine and Biology* 51, no. 8 (2006): 1971

M Besserve, J Martinerie, L Garnero "Improving quantification of functional networks with eeg inverse problem: Evidence from a decoding point of view." *NeuroImage* 55.4 (2011): 1536-1547.

Minkyu Ahn, Jun Hee Hong, Sung Chan Jun: "Feasibility of approaches combining sensor and source features in brain–computer interface." *Journal of Neuroscience Methods* 204 (2012): 168-178.

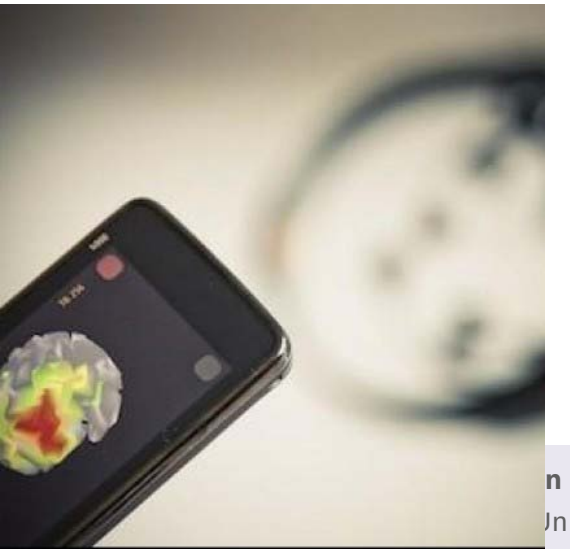
Smartphone Brain Scanner

Based on the Emotiv wireless transmission mechanism and either the EPOC head set or modified EasyCaps (thanks to Stefan Debener, Oldenburg)



First version ran on a Nokia platform

Version 2.0 works in generic Android platforms (Tested in Galaxy Note, Nexus 7,...)



https://github.com/SmartphoneBrainScanner

SBS functions current

netexplo Global observatory on digital society

Real time system

- Bayesian minimum norm 3D reconstruction with a variety of forward models (N=1024).
- Adaptive SNR model (β,σ) estimated every 10 sec.
- Update speed ~ 40 fps (Emotiv sample rate 128Hz, blocks of 8 samples)
- Selected frequency band option
- Spatial averaging in ”named” AAL regions

Mobile experiment set-ups, so far...

- Common spatial pattern- BCI
- Stimulus presentation options: image,text, audio
- Neuro-feedback

3D imaging to go ...
by source reconstruction on smartphones

Fig. 5: The 3D brain visualization application running on a Nexus 7 tablet. Blue color indicates cortical activity in the 8-12 Hz band.

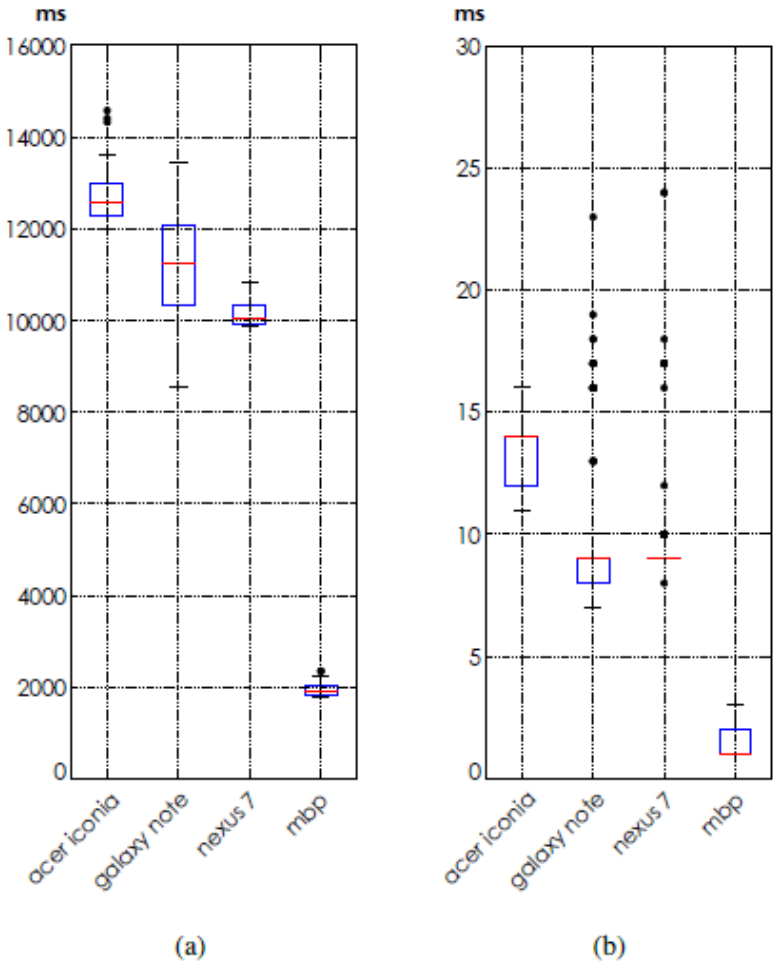
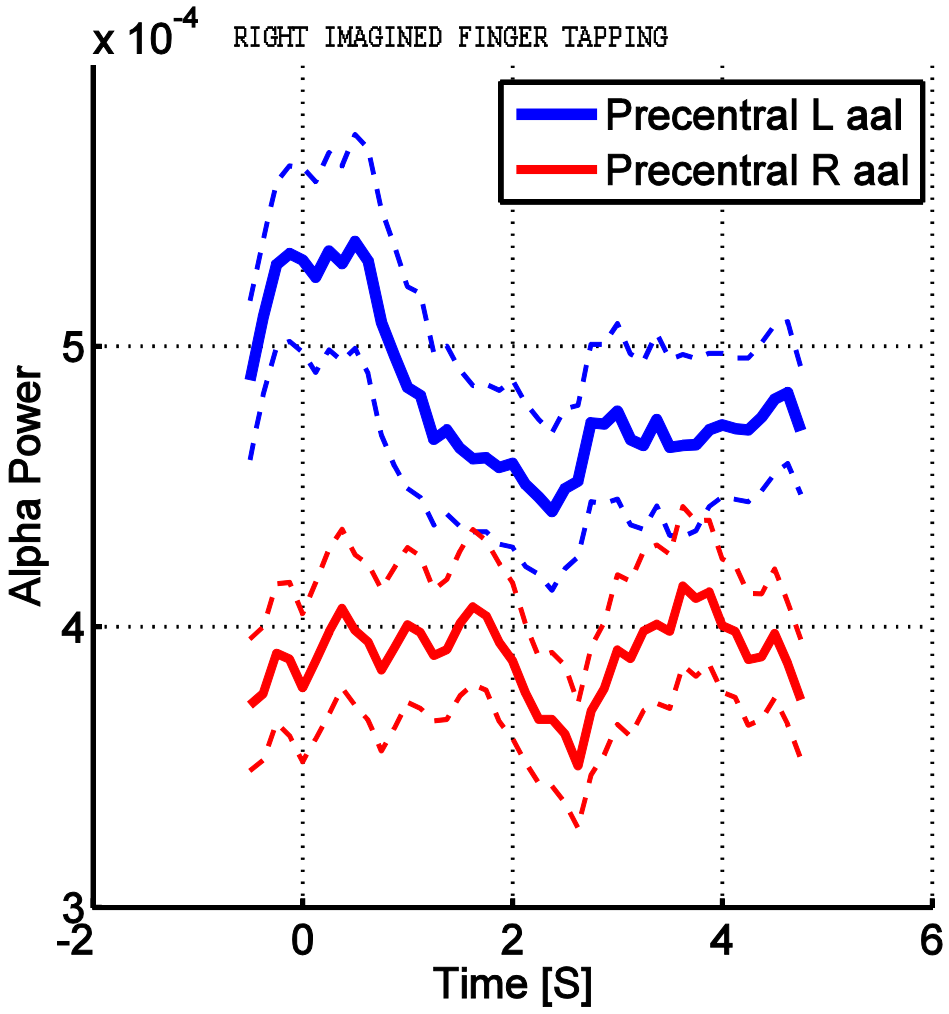


Fig. 4: Benchmark of 3D reconstruction performed on different devices. Time in ms needed for (a) model update and (b) time of the 3D reconstruction.

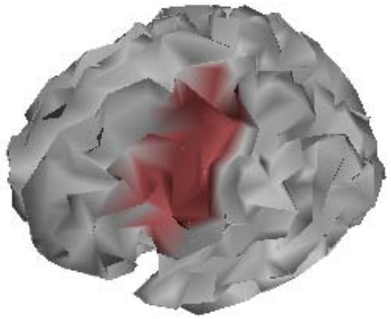
Do we get meaningful 3D reconstructions?

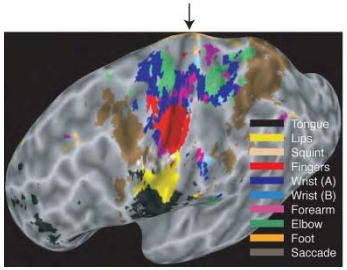


Imagined finger tapping

Left or right cued (at t=0)

Signal collected from an AAL region





Cimbi Center for integrated molecular brain imaging

Lars Kai Hansen

IMM, Technical University of Denmark

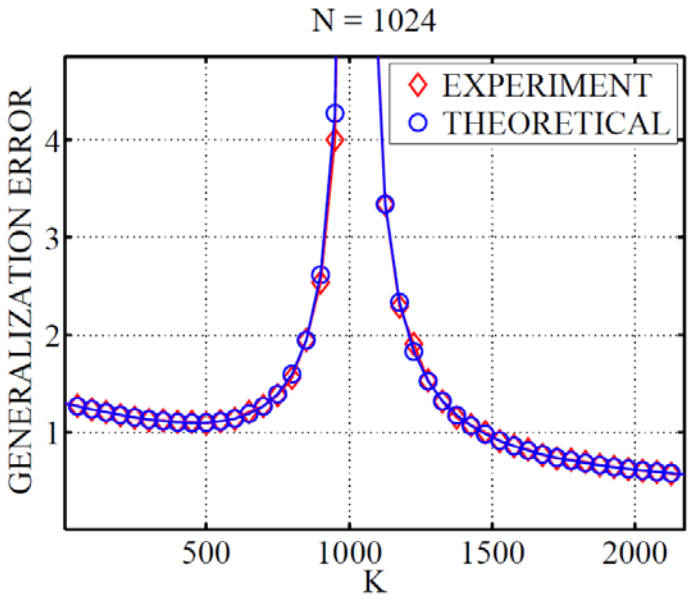
Meier, Jeffrey D., Tyson N. Aflalo, Sabine Kastner, and Michael SA Graziano. "Complex organization of human primary motor cortex: a high-resolution fMRI study." *Journal of neurophysiology* 100, no. 4 (2008): 1800-1812.

Intermezzo II: A pseudo-inverse ”in trouble...” exact results for an orthogonal (i.i.d. N(0,1)) A

$$Y_{k,t} = \sum_{n=1}^{N} A_{k,n} X_{n,t} + E_{k,t}.$$

$$||\mathbf{x}_0||^2 > \frac{N}{N-1}\sigma^2$$

$$K_{\min} = N - 1 - \sqrt{N(N-1)}\sqrt{\frac{\sigma^2}{||\mathbf{x}_0||^2}}$$



$$G(K) = \begin{cases} \left(1 - \frac{K}{N}\right)||\mathbf{x}_0||^2 + \frac{N-1}{N-K-1}\sigma^2 & K < N-1, \\ \infty & N-1 \le K \le N+1, \\ \frac{K-1}{K-N-1}\sigma^2 & K > N+1. \end{cases}$$

Hansen, L. K. Stochastic linear learning: Exact test and training error averages. Neural Networks 6(3), 393–396 (1993).

Barber, D., D. Saad, and P. Sollich. Finite-size effects and optimal test set size in linear perceptrons. J Phys A 28 1325-34 (1995).

Why sparse reconstruction?

Neuro-physiological

- Evidence that brain ”modes” are dipolar (Delorme et al., 2012)
- Brain modules are specialized (From functional localization hypothesis -> networks of interacting modules – sparse networks)

Modeling

- Imaging problem is severely ill-posed (Pascual-Marqui et al. 2002)
- Sparsity promoting priors can improve uniqueness of solutions
- Sparsity may be in a basis set (Haufe et al. 2011).

A Delorme, J. Palmer, J. Onton, R. Oostenveld, S. Makeig. Independent EEG sources are dipolar. PloSone, 7 (2) e30135, 2012.

RD Pascual-Marqui, M. Esslen, K. Kochi, D. Lehmann, et al., Functional imaging with low-resolution brain electromagnetic tomography (loreta): a review,” Methods and findings in experimental and clinical pharmacology, 24(C):91–95, 2002.

S. Haufe, R. Tomioka,T. Dickhaus, C. Sannelli, B Blankertz, G, Nolte, KR Müller. Large-scale EEG/MEG source localization with spatial flexibility. NeuroImage, 54(2), 851-859. (2011)

Sparsity promoting priors

$$Y_{k,t} = \sum_{n=1}^{N} A_{k,n} X_{n,t} + E_{k,t}$$

Direct search for sparse solutions (X)

Feature selection /pruning, active sets

Automatic relevance determination,

Spike and slab,

Variational Garrote

Convex relaxations

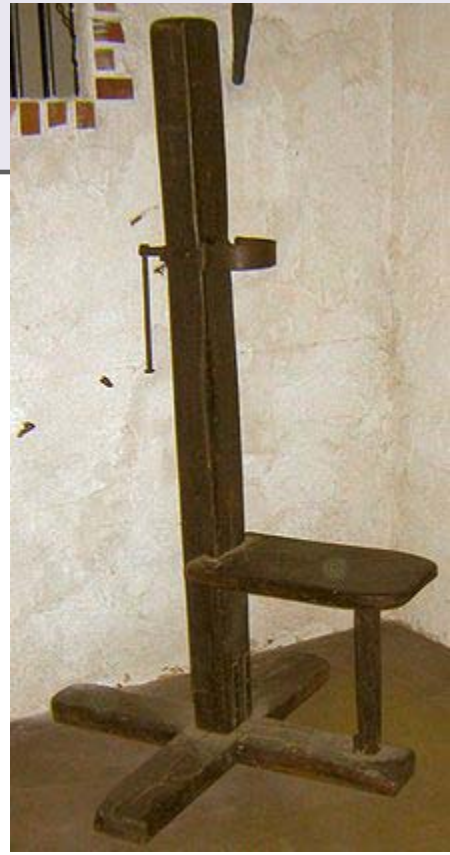
$||X||_1$ and other regularizers,

Least angle regression, homotopy, path methods

Variational Garrote (Kappen, 2011)

$$Y_{k,t} = \sum_{n=1}^{N} A_{k,n} X_{n,t} + E_{k,t}.$$

$$Y_{k,t} = \sum_{n=1}^{N} A_{k,n} S_{n,t} X_{n,t} + E_{k,t}$$



Introduce binary indicators for the support of the solution

- Inspired (name..) by Breiman's "non-negative garrote"
- Similar to spike and slap, or Bernoulli-Gauss priors
- Variational inference

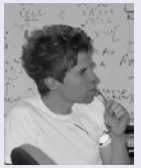
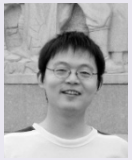
EEG: Potential separation of time scales (s variables are smooth in time)

L. Breiman. "Better subset regression using the nonnegative garrote". Technometrics, 37(4):373–384, 1995.

HJ Kappen. "The variational garrote." *arXiv preprint arXiv:1109.0486,* 2011.

M Titsias, M Lazaro-Gredilla. "Spike and Slab Variational Inference for Multi-Task and Multiple Kernel Learning". Advances in Neural Information Processing Systems 24, 2339–2347, 2011.

Multiple Measurement Vectors (MMV)


M-SBL (Wipf & Rao, 2007) models repeated measurement, hence, a source is present or absent for the whole time frame. Each source has a Tikhonov $||X||_2$ regularizer with estimated strength (ARD, SBL "Sparse Bayesian Learning")

T-MSBL (Zhang & Rao, 2011) also assumes a block-structure so that each source is a block. Temporal correlations are modeled in the blocks and ARD control parameter identifies active sources (blocks) . Extensive simulations show state of the art performance.

Yet, to compare to Ziniel et al. (2010), who use approximate BP to model both smoothness in both amplitude and support with spike and slab prior.

Zhilin Zhang, B.D. Rao. "Sparse Signal Recovery with Temporally Correlated Source Vectors Using Sparse Bayesian Learning". IEEE Journal of Selected Topics in Signal Processing, 5(5):912-926, 2011.

D. P. Wipf, B.D. Rao. "An empirical Bayesian strategy for solving the simultaneous sparse approximation problem," *IEEE Trans. on Signal Processing,* 55(7):3704–3716, 2007.

J. Ziniel, LC. Potter, P. Schniter. "Tracking and smoothing of time varying sparse signals via approximate belief propagation," in *Proc. Of the 44th Asilomar Conference on Signals, Systems and Computers,* 808–812 (2010).

Variational inference

$$Y_{k,t} = \sum_{n=1}^{N} A_{k,n} S_{n,t} X_{n,t} + E_{k,t}$$

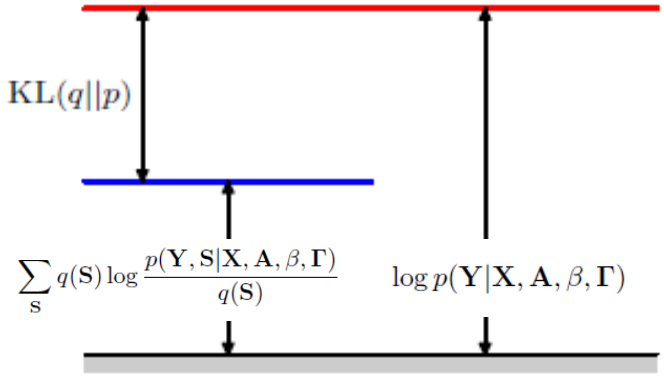
The likelihood function is given by

$$p(\mathbf{Y}|\mathbf{X}, \mathbf{A}, \beta, \mathbf{\Gamma}) = \sum_{\mathbf{S}} p(\mathbf{Y}|\mathbf{X}, \mathbf{A}, \mathbf{S}, \beta) p(\mathbf{S}|\mathbf{\Gamma})$$

$$\log p(\mathbf{Y}|\mathbf{X}, \mathbf{A}, \beta, \mathbf{\Gamma}) = \sum_{\mathbf{S}} q(\mathbf{S}) \log \frac{p(\mathbf{Y}, \mathbf{S}|\mathbf{X}, \mathbf{A}, \beta, \mathbf{\Gamma})}{q(\mathbf{S})} + \sum_{\mathbf{S}} q(\mathbf{S}) \log \frac{q(\mathbf{S})}{p(\mathbf{S}|\mathbf{Y}, \mathbf{X}, \beta, \mathbf{\Gamma})}$$

$$\log p(\mathbf{Y}|\mathbf{X}, \mathbf{A}, \beta, \mathbf{\Gamma}) > -F(\mathbf{X}, \beta, \mathbf{\Gamma}, \mathbf{M})$$

$$\equiv \sum_{\mathbf{S}} q(\mathbf{S}|\mathbf{M}) \log \frac{p(\mathbf{Y}, \mathbf{S}|\mathbf{X}, \mathbf{A}, \beta, \mathbf{\Gamma})}{q(\mathbf{S}|\mathbf{M})}$$



Variational inference w./ fully factored q(S)

$$F(\mathbf{X}, \beta, \mathbf{\Gamma}, \mathbf{M}) =$$

$$-\sum_{\mathbf{S}} q(\mathbf{S}|\mathbf{M}) \left[\log p(\mathbf{Y}|\mathbf{S}, \mathbf{X}, \mathbf{A}, \beta, \mathbf{\Gamma}) + \log p(\mathbf{S}|\mathbf{\Gamma}) - \log q(\mathbf{S}|\mathbf{M})\right]$$

$$q(\mathbf{S}) = \prod_{n,t} \left[M_{n,t} S_{n,t} + (1 - M_{n,t})(1 - S_{n,t})\right]$$

$$F(\mathbf{X}, \beta, \mathbf{\Gamma}, \mathbf{M}) = -\frac{KT}{2} \log \frac{\beta}{2\pi} + \frac{\beta}{2} \sum_{k,t} (Y_{k,t} - \sum_{n} A_{k,n} M_{n,t} X_{n,t})^2$$

$$+ \sum_{n,t} M_{n,t}(1 - M_{n,t}) X_{n,t}^2 \sum_{k} A_{k,n}^2$$

$$- \langle \log p(\mathbf{S}|\mathbf{\Gamma}) \rangle_q + \sum_{n,t} M_{n,t} \log M_{n,t} + (1 - M_{n,t}) \log(1 - M_{n,t})$$

Variational inference w./ fully factored q(S)

HJ Kappen. "The variational garrote." *arXiv preprint arXiv:1109.0486,* 2011.

VG: factorized prior

$$p(\gamma) = \frac{1}{1+e^{-\gamma}}$$

$$p(\mathbf{S}) = \prod_{n,t=1}^{N,T} p(S_{n,t}) = \prod_{n,t=1}^{N,T} (p(\gamma)S_{n,t} + (1-p(\gamma))(1-S_{n,t}))$$

Streaming prior, first order Markov chain

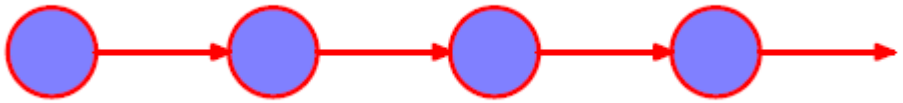
$$p(S_{n,t}|S_{n,t-1}) = \mathbf{\Gamma}_{S_{n,t},S_{n,t-1}}$$

$$p(\mathbf{S}) = p(S_{n,1}) \prod_{n,t=2}^{N,T} p(S_{n,t}|S_{n,t-1})$$

Temporally smooth, sparse prior

$$p(S_{n,t}|S_{n,t-1}) = \Gamma_{S_{n,t},S_{n,t-1}}$$

$$p(\mathbf{S}) = p(S_{n,1}) \prod_{n,t=2}^{N,T} p(S_{n,t}|S_{n,t-1})$$



Stationary distribution

$$\boldsymbol{\pi}_* = \boldsymbol{\Gamma}\boldsymbol{\pi}_*$$

Sparsity

$$\frac{\boldsymbol{\pi}_*(1)}{\boldsymbol{\pi}_*(0)} = \frac{\boldsymbol{\Gamma}_{1,0}}{\boldsymbol{\Gamma}_{0,1}}$$

Temporal smoothness

$$\Gamma_{0,0} = p(s_t = 0|s_{t-1} = 0)$$

Update rules for smooth, sparse prior

$$\mathbf{B} = \mathbf{A}^T\mathbf{Y}/K \qquad \chi^0 = \mathbf{A}^T\mathbf{A}/K$$

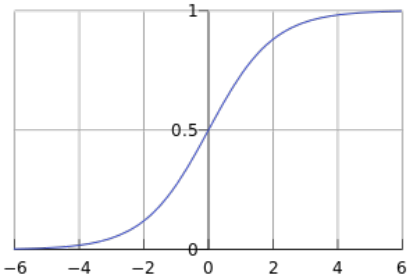
$$\chi_t = \chi_{n,n',t} = \chi^0_{n,n'} M_{n,t} + (1 - M_{n,t})\chi^0_{n,n}\delta_{n,n'}$$

$$\mathbf{X}_t = \chi_t^{-1}\mathbf{B}_t$$

$$M_{n,t} = \sigma\left(g_1 + \frac{g_2}{2}(M_{n,t+1} + M_{n,t-1}) + \frac{\beta K}{2} X^2_{n,t}\chi_{n,n}\right)$$

$$\text{with } g_1 = \log\left(\frac{\Gamma_{1,0}\Gamma_{0,1}}{\Gamma_{0,0}}\right), g_2 = \log\left(\frac{\Gamma_{1,1}\Gamma_{0,0}}{\Gamma_{1,0}\Gamma_{0,1}}\right).$$

$$\beta^{-1} = \left(\sum_{k,t} Y^2_{k,t} - \sum_{n,t} M_{n,t}X_{n,t}B_{n,t}\right)/KT$$



Evaluation

Single time shot:

VG direct space Vs. VG dual Vs. ARD Vs

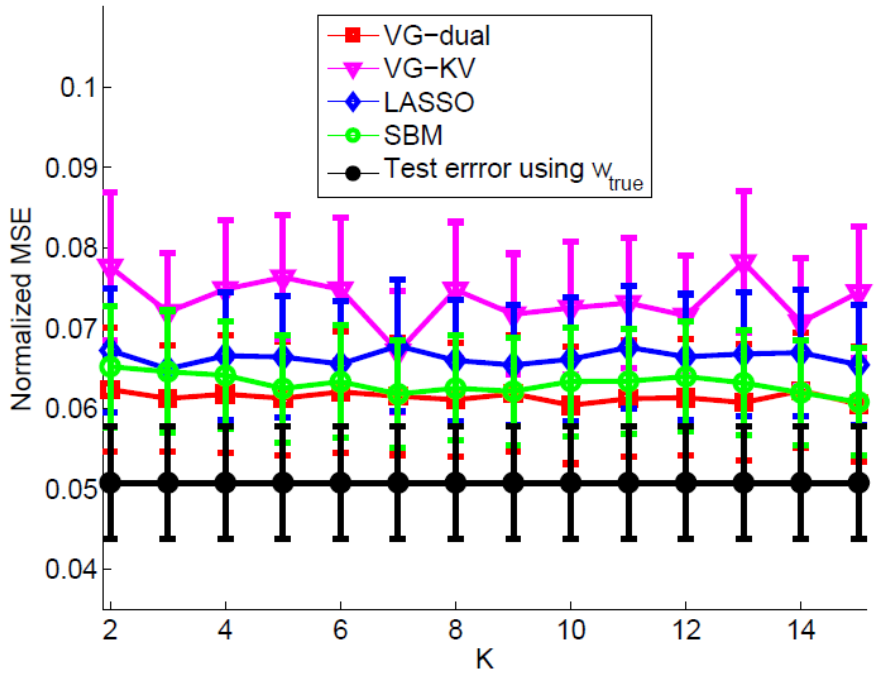
Lasso, L1 convex relaxation

Temporal case: Vs. Multiple Measurement Vectors (MMV)

When A almost orthogonal (A~N(0,1))

When A is an EEG head model (”Emocap”)

Single time shot / EEG head model from SPM (N=8196, K=128)



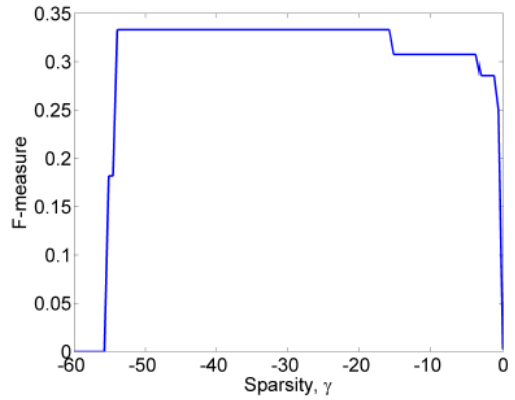


Fig. 5. Ability to retrieve the planted sources (F) as function of sparsity control γ.

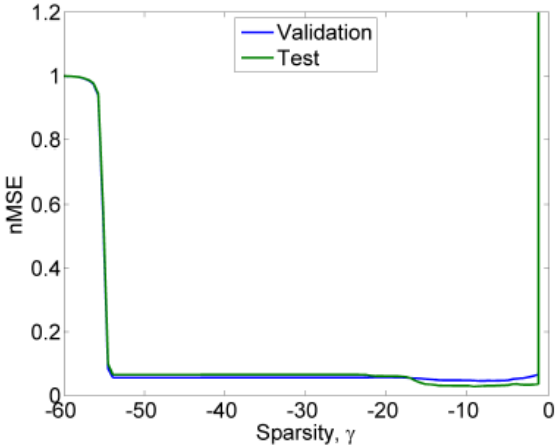


Fig. 6. Cross-validation error as function of γ.

Two split cross-val: test(10), train(118)

Within train K-fold cross val to determine γ

Single time shot

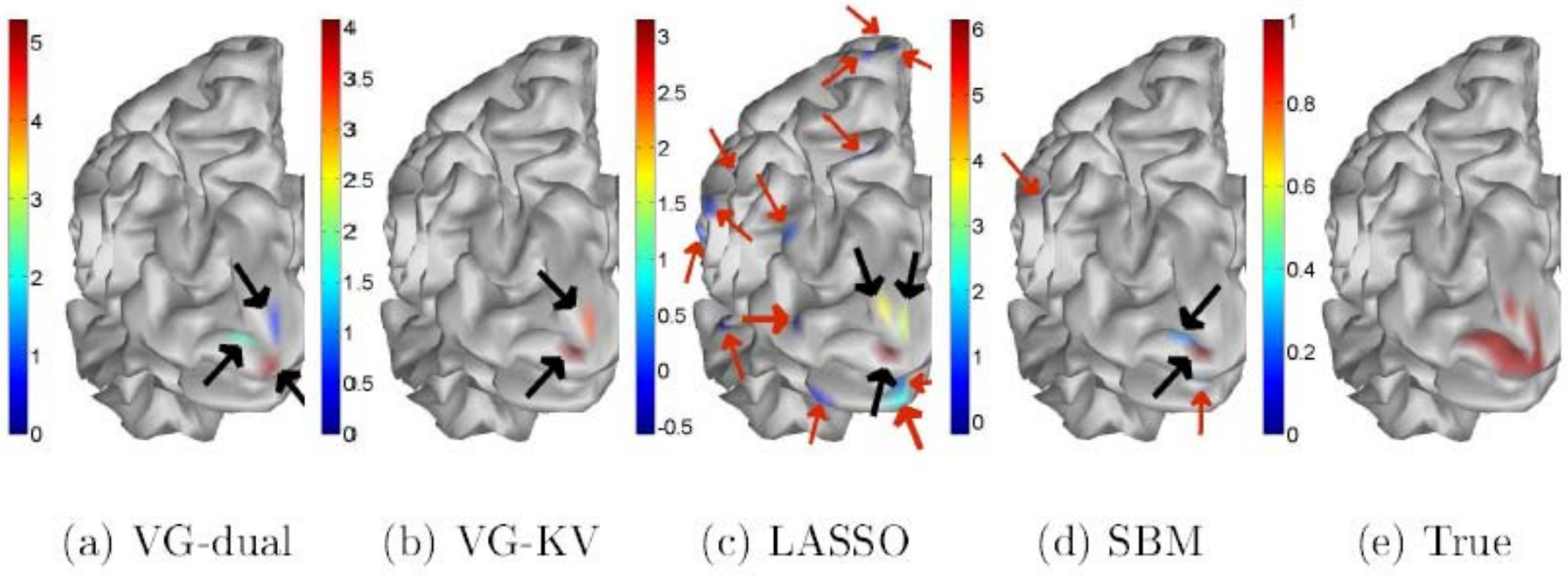
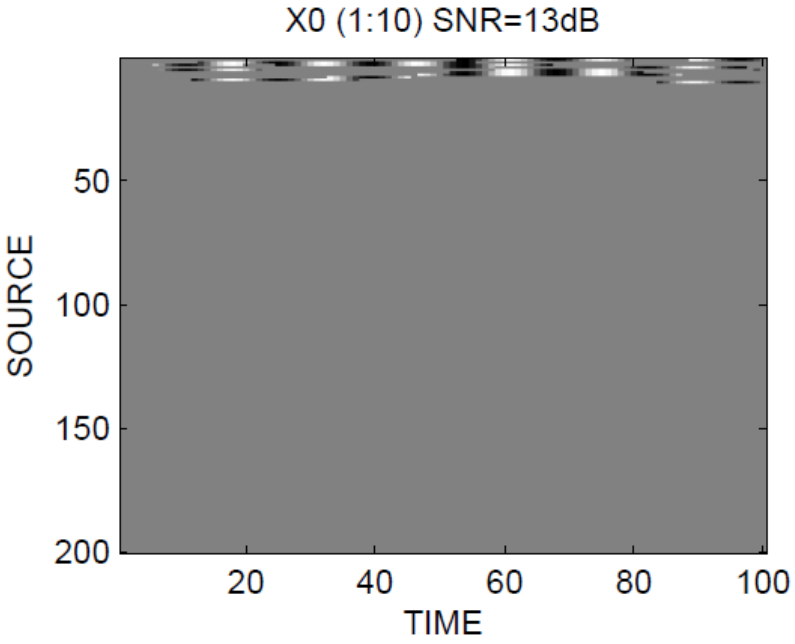
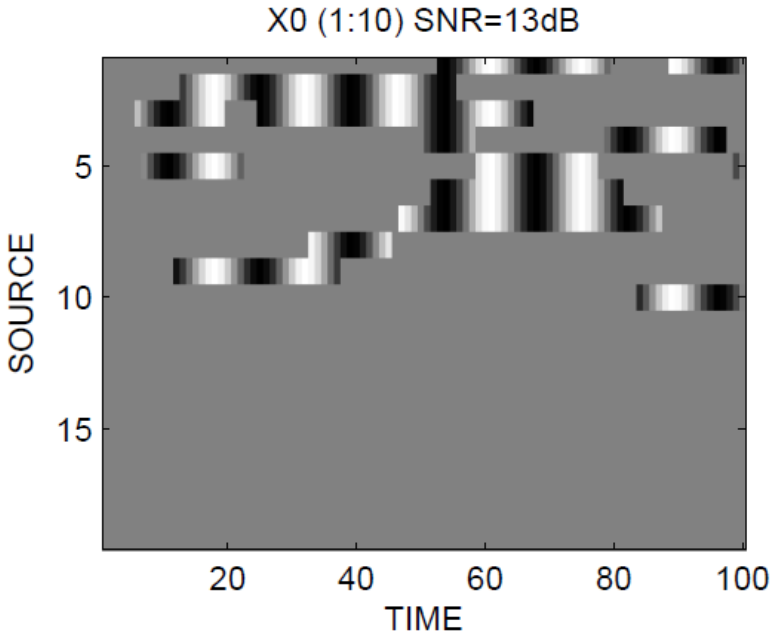


Figure 4.5: Sources estimated in the context of a 3D cortex structure are compared with the 'true' distribution. The algorithms are optimized through ten-fold cross-validation with respect to one parameter; for the VGs the sparsity level γ, for LASSO the regularization parameter λ, and for SBM the precision of the noise β. The solutions for the VGs correspond to $\mathbf{v}$ including a threshold on the activation $\mathbf{m}$ set to 0.5. For LASSO and SBM the solution presented are $\mathbf{w}$ with a treshold on the weights set to 10^{-10}. Heavy or thin arrows indicate sources with magnitudes larger or less than 0.5, respectively. Black arrows indicate true sources and red false sources. View is from the back of the left hemisphere. No sources are found in the right hemisphere for the VGs, only low-strength sources for LASSO and one low strength for SBM. Note individual color maps are used.

Temporal simulation, A orthogonal (N=200, K=30, T=100)

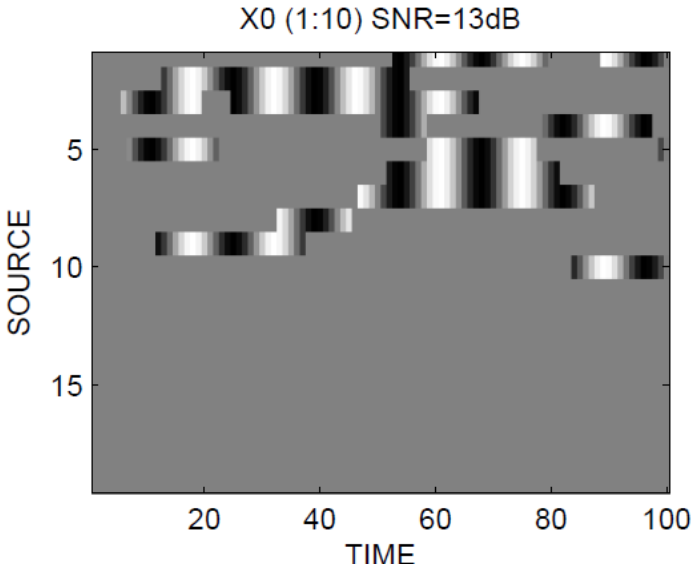


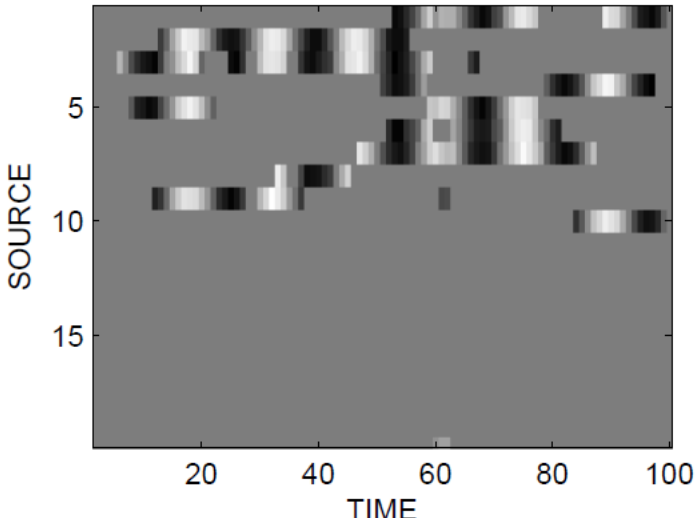


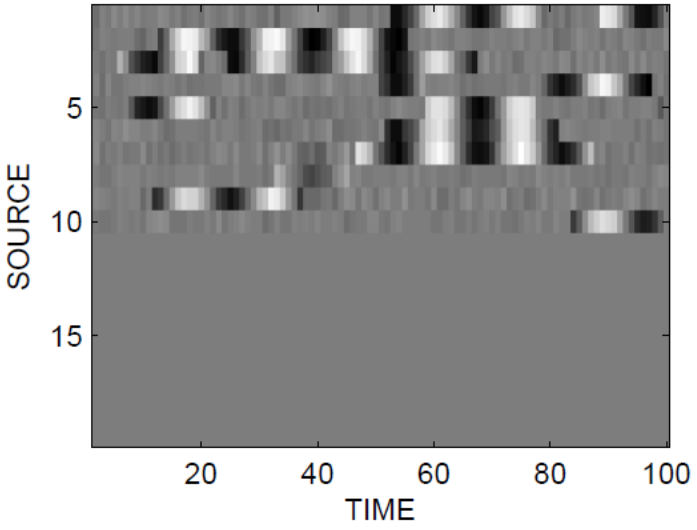
Active sources n=1:10

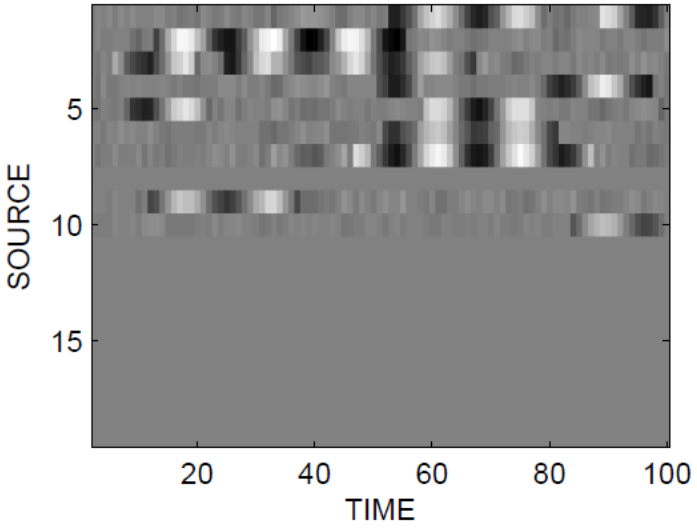
Sine wave time function in X

Temporal simulation, A orthogonal (N=200, K=30, T=100)



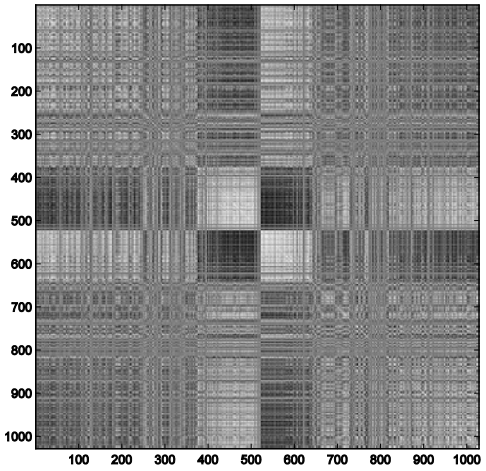




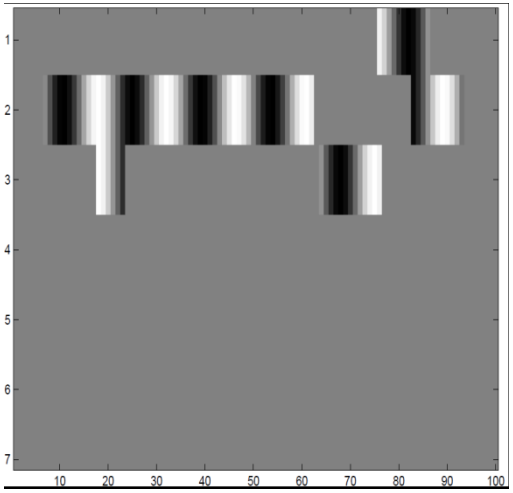


Temporal simulation, A is an EEG forward model (N=1028, K=14, T=100)

$\mathbf{A}^T\mathbf{A}$



$N_{active}=3, \quad s.*x_0$



100 repetitions

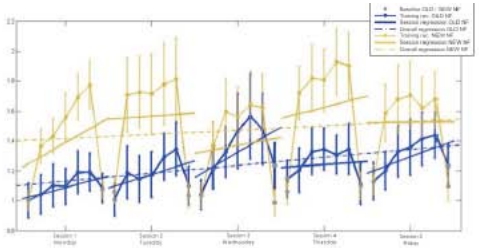
	VG	TMSBL	MMV
Mean Acc	0.9992± 0.0004	0.9906± 0.0072	0.9888± 0.0292

All dressed up – where to go?

Low cost equipment => Large scale social neuroscience...

Sensible DTU experiment tracking a unique population of 135 students with smartphones

Neuro-feedback in near-natural conditions (CF. Jensen et al, ”Training your brain on a tablet”)



Cimbi Center for integrated molecular brain imaging

Lars Kai Hansen

IMM, Technical University of Denmark

DTU

Conclusion

Sparse source reconstruction may provide us with real-time 3D EEG imaging

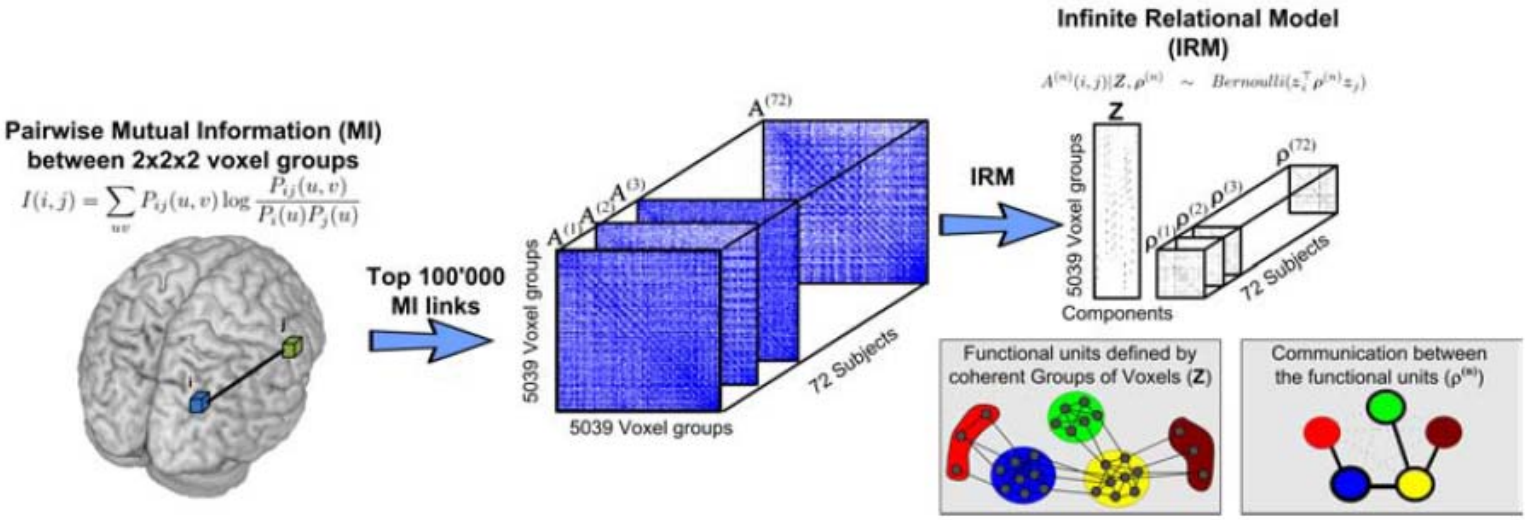
Separation of time scales using ”spike and slab” like representation

Variational inference allows kernel trick like dual formulation with linear scaling pr update

Promising results on realistic, coherent forward model

Detecting networks with relational models

Hypothesis: Networks variability in fMRI resting state fluctuations separates a group of MS patients from normal group (Ntot =72, 40MS + 32HC)

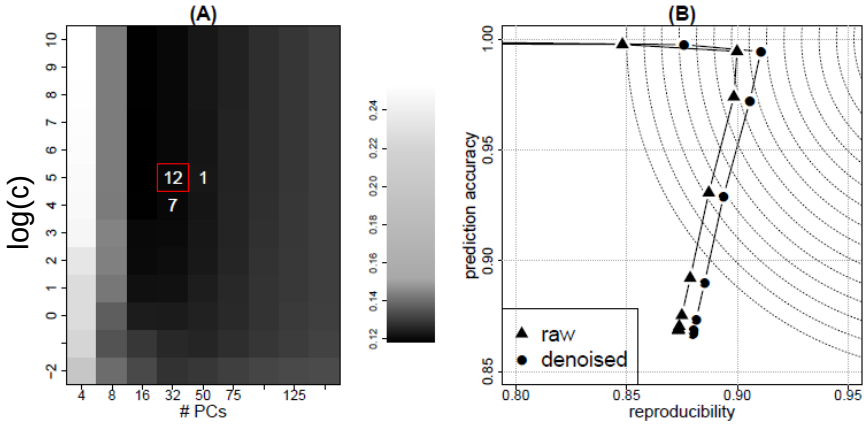


Basic measure: Mutual information between time series (can detect similarity by modulation)

The Bayesian non-parametric Infinite Relational Model (IRM) provides functional network segregation (communities) and a summary of the intra- and inter-community communication patterns

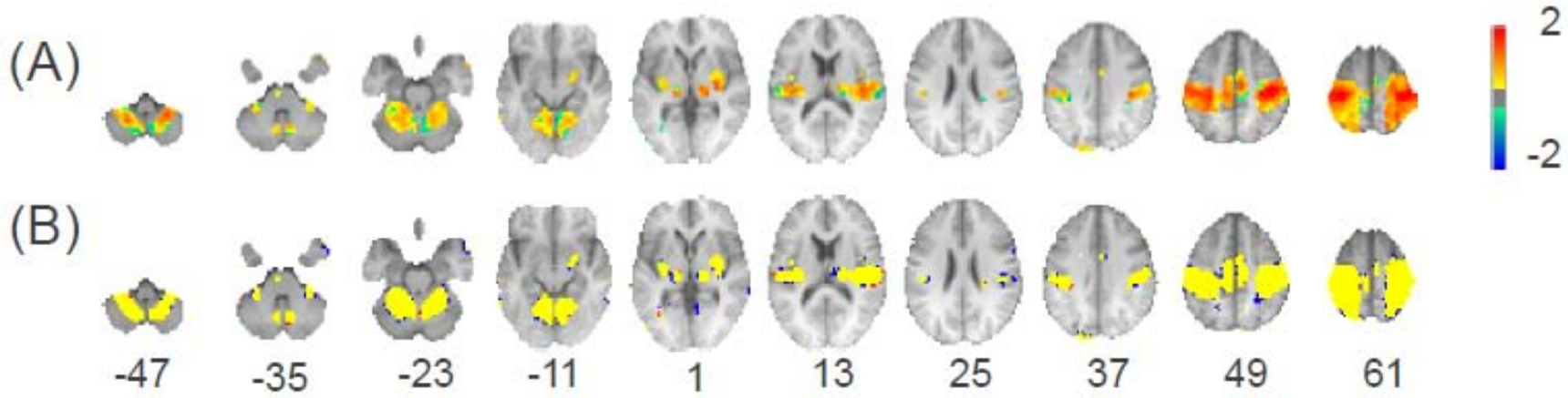
M Mørup, KH Madsen, AM Dogonowski, H Siebner, LK Hansen. Infinite relational modeling of functional connectivity in resting state fMRI. NIPS 23 (2010).

Denoising by kernel PCA helps fMRI decoding...



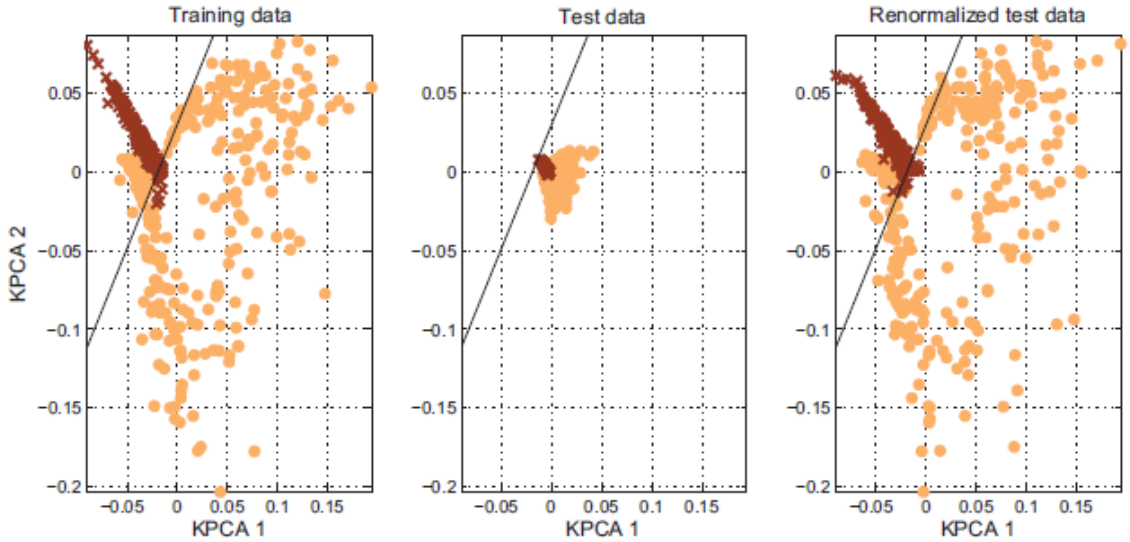
(A) NPAIRS (w. Stephen Strother)
Comparison of resampling z-score = z-kPCA – z-Raw

(B) FDR corrected:
yellow: consensus,
blue: only kPCA,
red: only raw

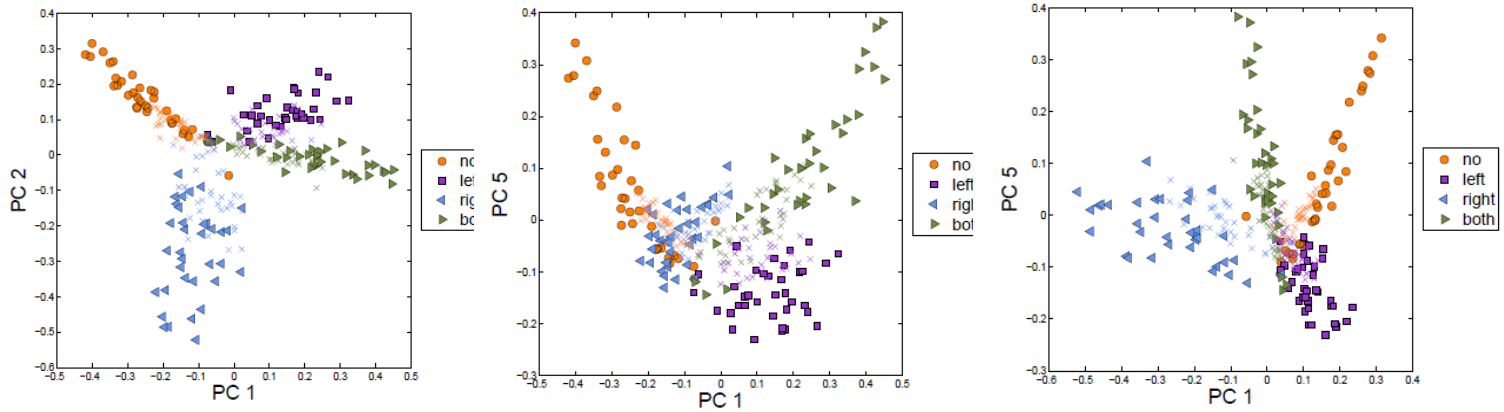


PM Rasmussen, TJ Abrahamsen, KH Madsen, LK Hansen: Nonlinear denoising and analysis of neuroimages with kernel principal component analysis and pre-imageestimation, NeuroImage 60(3):1807-1818 (2012).

Variance inflation in ill-posed (kernel) factor models



XOR fMRI: D= 75,257 N=576



T.J. Abrahamsen and L.K. Hansen. A Cure for Variance Inflation in High Dimensional Kernel Principal Component Analysis. Journal of Machine Learning Research 12:2027-2044 (2011).

Acknowledgments

Lundbeck Foundation (www.cimbi.org, Stahlhut)
Danish Research Councils

www.imm.dtu.dk/~lkh
hendrix.imm.dtu.dk