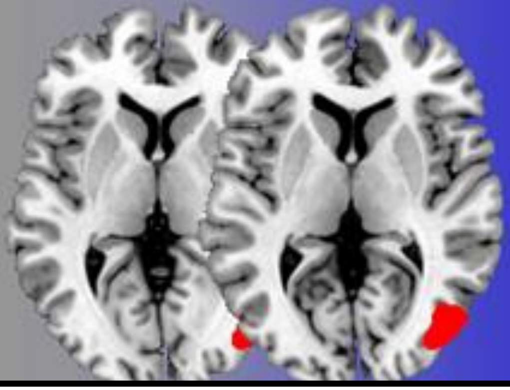
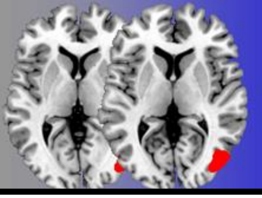
EMORY
ROLLINS SCHOOL OF PUBLIC HEALTH

CBIS
Center for Biomedical Imaging Statistics



BIOS 516

Statistical Analysis of Neuroimaging Data

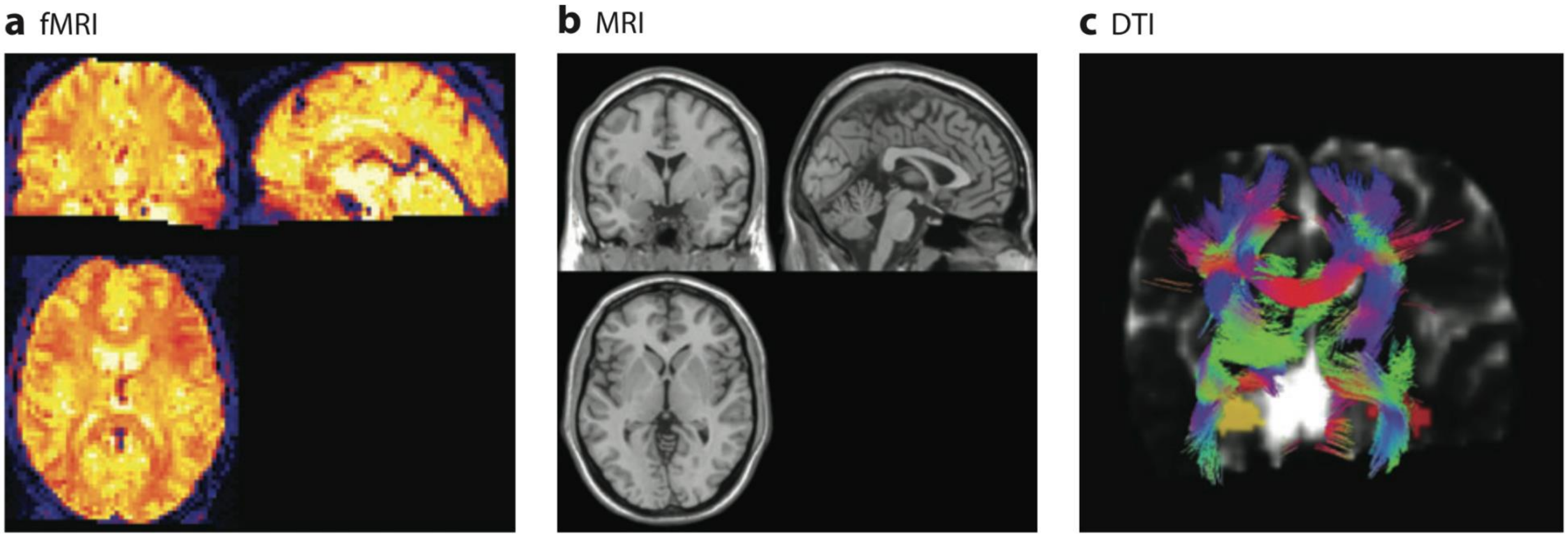


Acknowledgments:

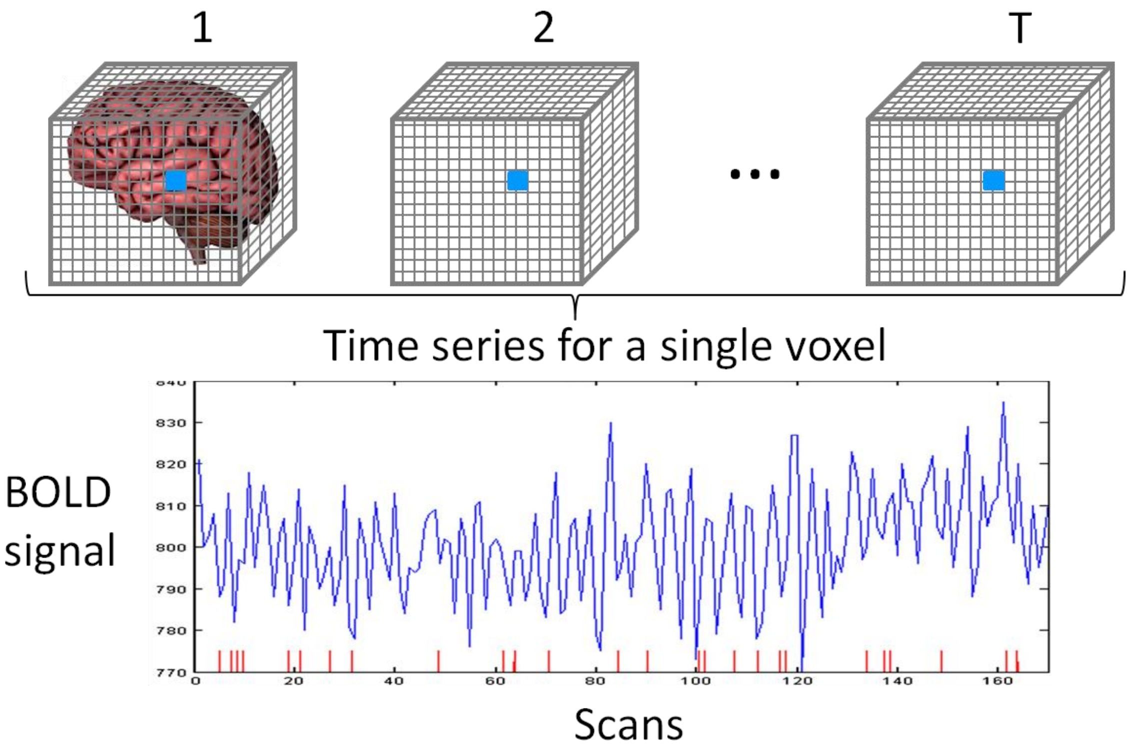
Most of these slides were taken from a presentation by Dr. Ying Guo.
Thanks Ying!

Review from last time…

- Structural Imaging modalities
  - T1 and T2 weighted MRI, CT, dMRI (DTI)
- Functional Imaging modalities
  - **fMRI,** PET, MEG & EEG



fMRI data



- Each fMRI image consists of ~100K voxels
- During the course of the experiment, hundred of images are acquired (~ one every 2 sec)
- Multiband (simultaneous multislice) decreases to 0.5-1 sec (see end of slides)
- One voxel → one BOLD time series

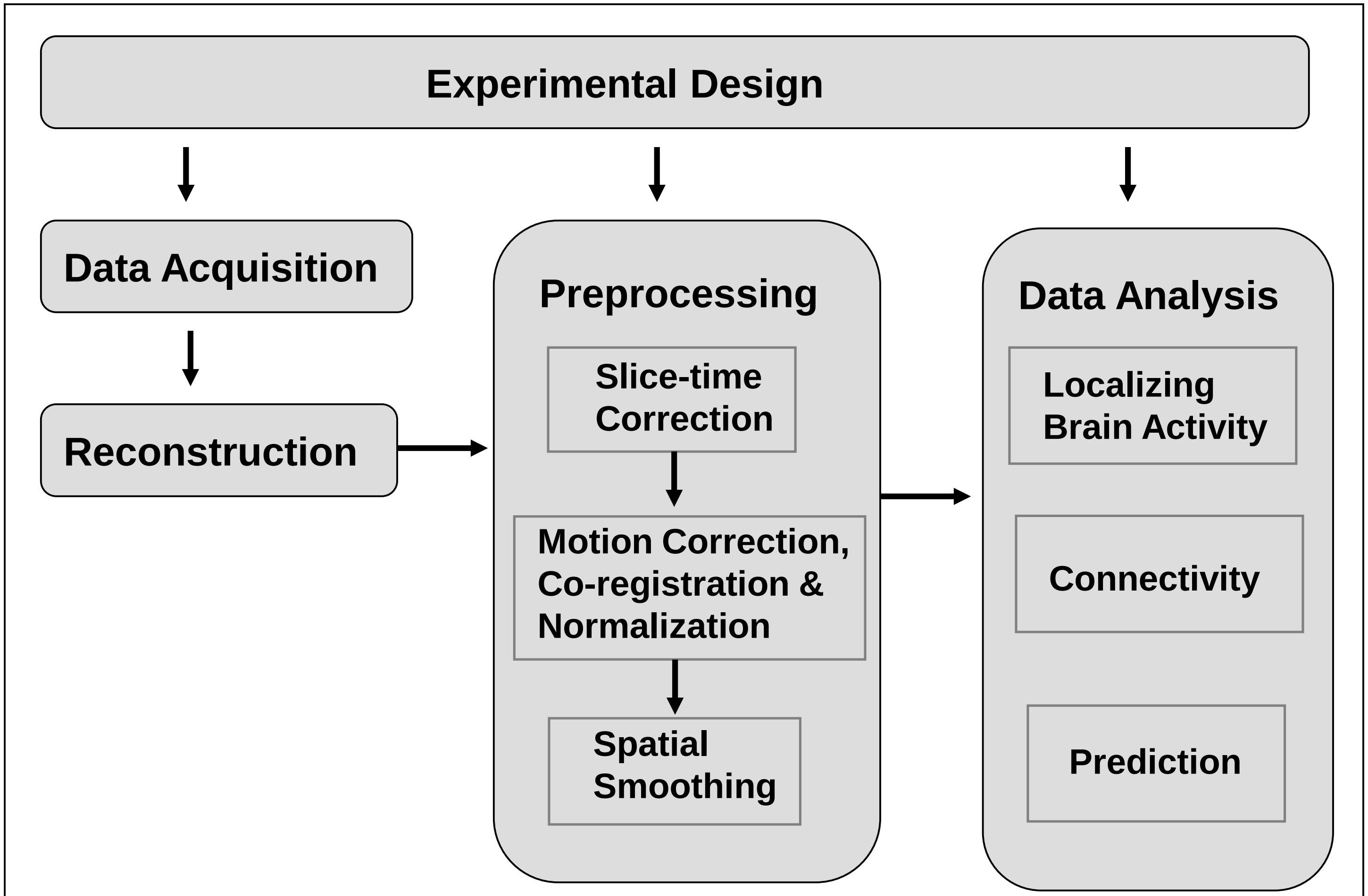
There are multiple goals in the statistical analysis of neuroimaging data:

- **ACTIVATION**: Localizing brain areas activated by the experimental task (Brain Mapping)
- **BRAIN CONNECTIVITY and NETWORK ANALYSIS**
- **PREDICTION**: making predictions about psychological or disease states

Challenges

- The statistical analysis of fMRI data is challenging.
  - It is a massive data problem
  - The signal of interest is relatively weak (only 0.5-3% change in intensity)
  - The data exhibits complex temporal and spatial structure

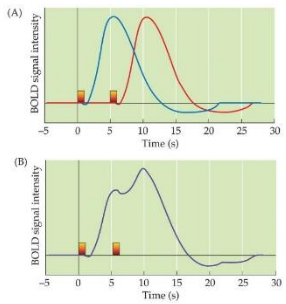
Analysis Procedure

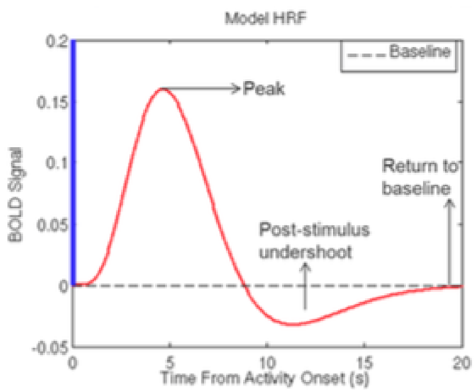


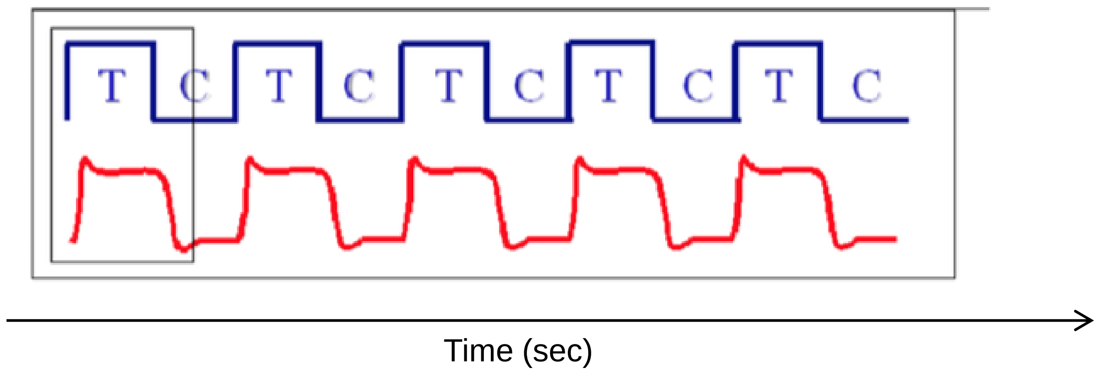
Credits: Martin Lindquist, Johns Hopkins University

fMRI study designs

- **Block Designs**: stimuli of the same condition are grouped together in blocks



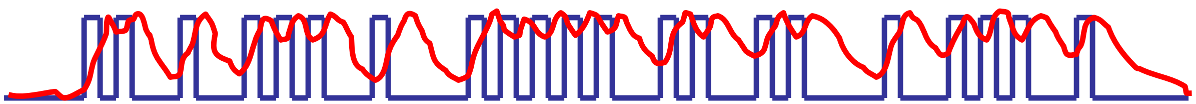




- **PRO**: Repeating the stimulus in a block causes a large total signal change – increases statistical power to detect activation
- **CON**: Can't directly estimate features of the HRF

fMRI study designs

- **Event-related Designs**: Allow different stimuli to be presented in arbitrary sequences, allows randomization of conditions.



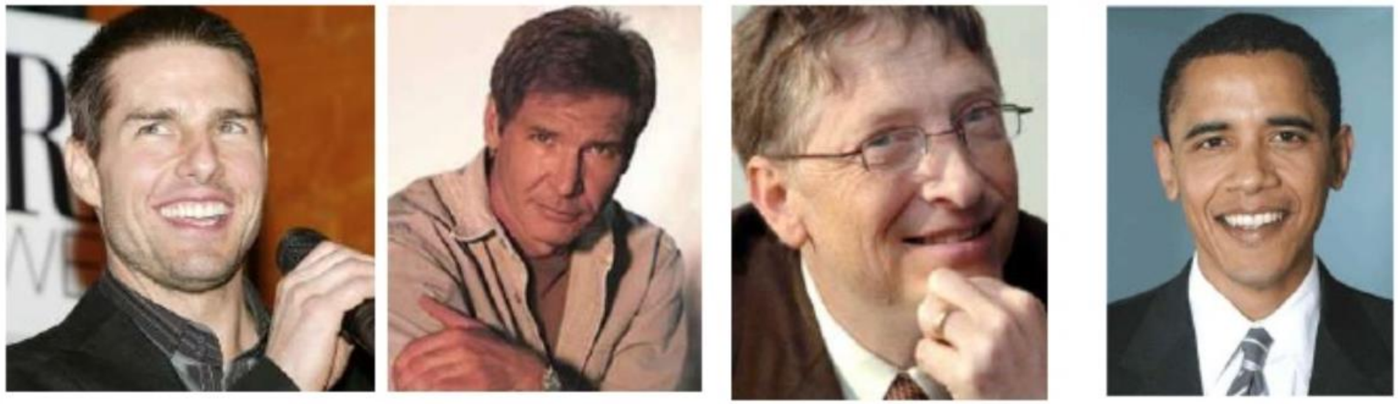
Figures: from Amaro and Barker, 2006 and lcni.uoregon.edu/~ray/

- **PRO**: Can precisely observe the actual HRF – thus allowing for the estimation of features of the HRF
- **CON**: Reduced statistical power to detect BOLD differences between different conditions. Lower signal change, may be <1%. For block design, 3-5%.

Simple task example

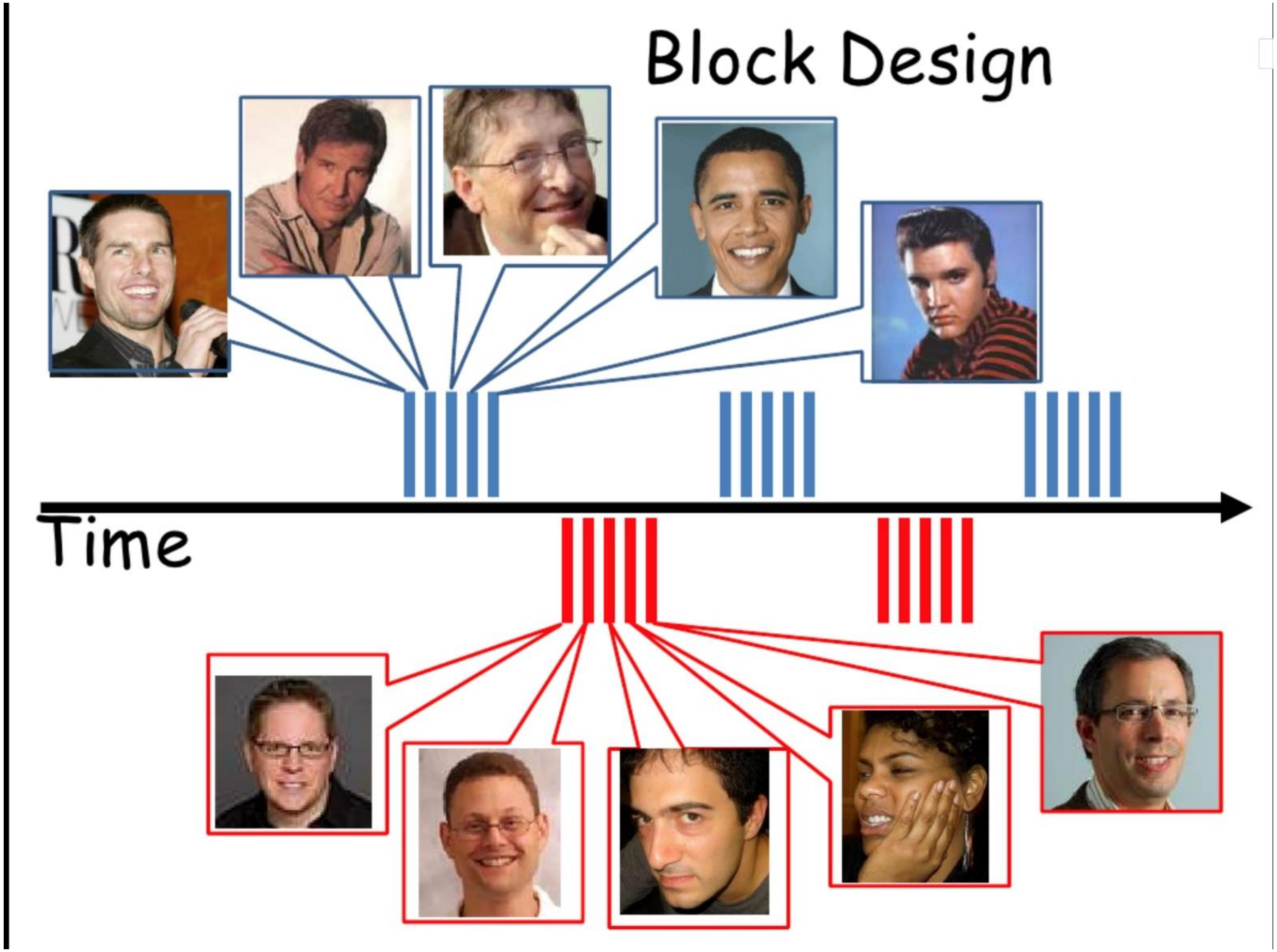
Facial recognition task

Task A: View famous faces

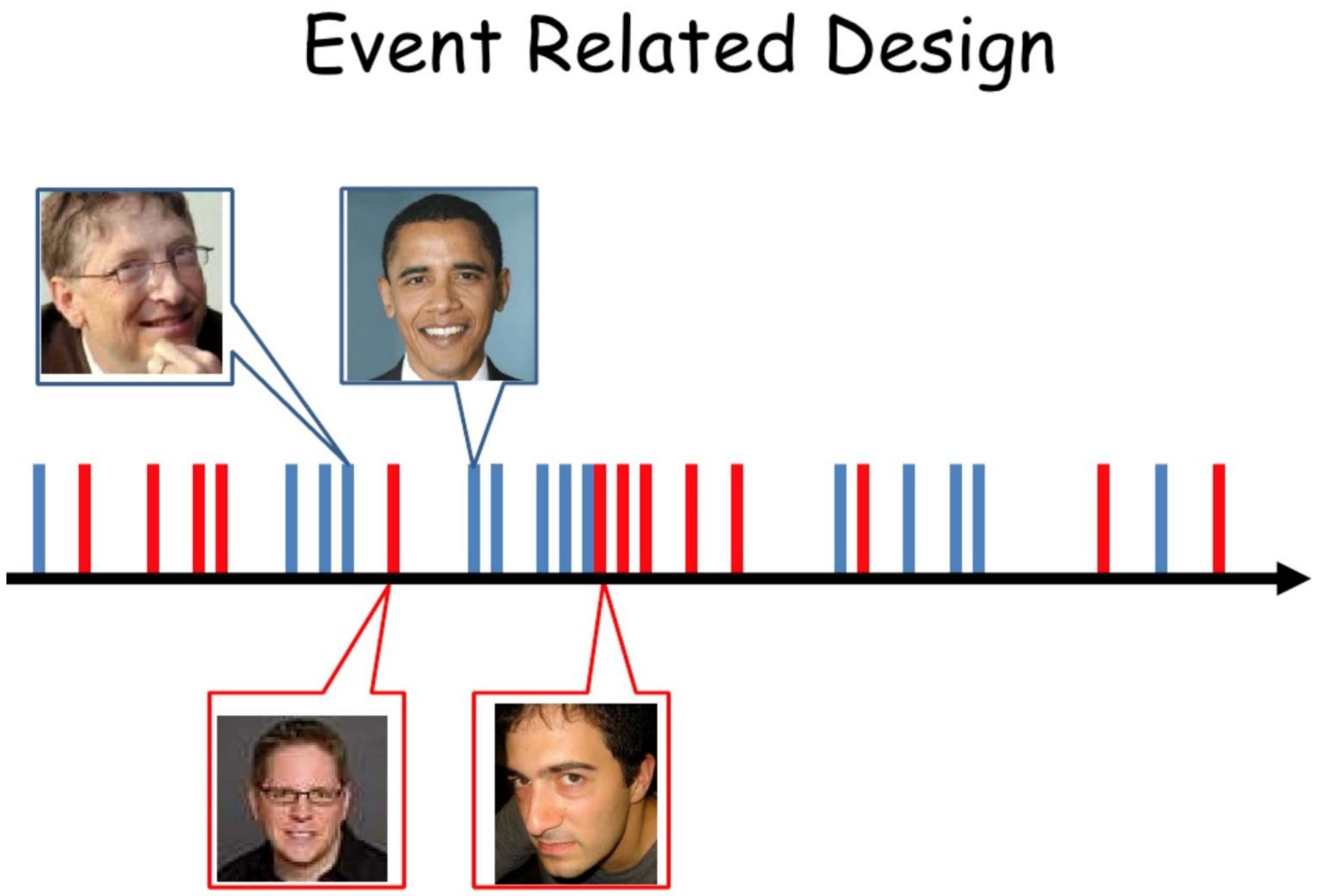


Task B: View non-famous faces

Simple task example



Simple task example




BOLD response in
block v. event related (slow)

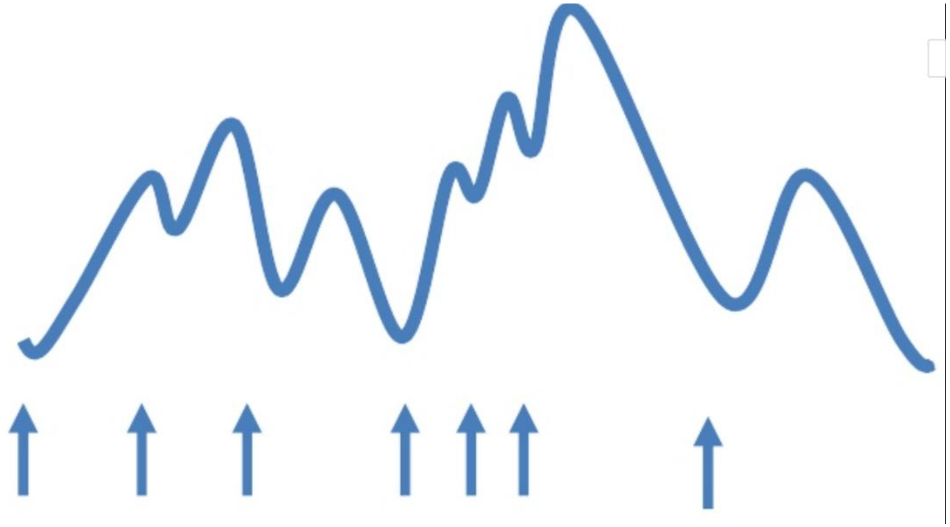
Block Design Issues

- Repetitions can get predictable, reducing activation

- Timing Issues:
  - Ideally 15-20 sec on, then 15-20 sec off
  - Long enough for HRF to relax in between presentations
  - Short enough for many comparison blocks within short time

Event-Related Design Issues

- Slow: Waiting 12+ seconds in between each event to allow HRF to relax is inefficient
- Gap spacing >4 seconds to avoid HRF blurring
- Jitter spacing to record different parts of the HRF and avoid correlation with other functions like heart rate and breathing



Task-related activity

- How to capture task-related activity in a noisy brain? Use cognitive subtraction/contrasts (task vs. control)

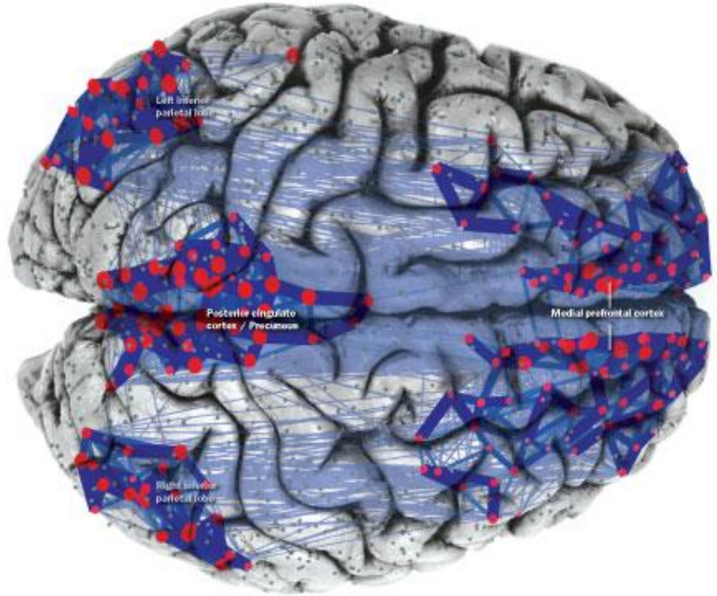
Image of task A

Image of task B

Image of task A- Image of task B

- **Resting-state fMRI studies:**
  - **No task**/stimulus
  - Acquire scans while subjects are left to think for themselves
  - May reflect a natural or more common mode of neural processing

Brain is **not** silent at rest!

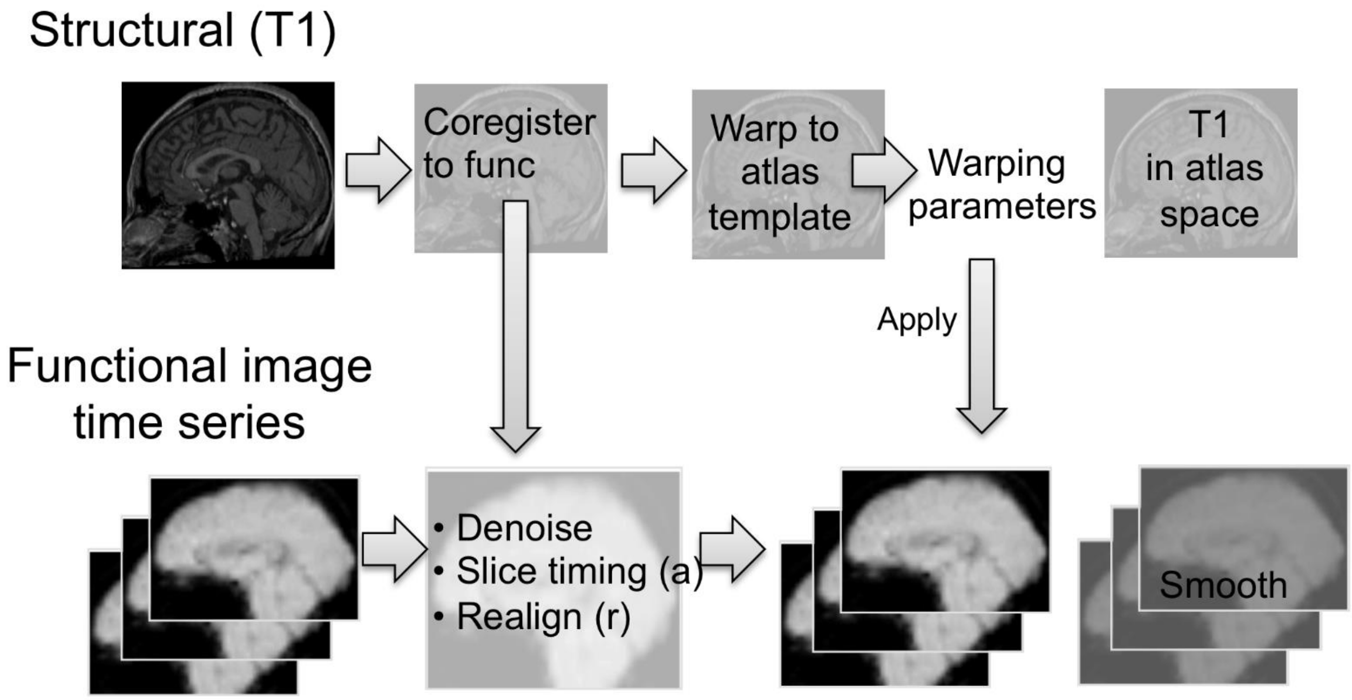


Default Mode Network

Sources of Noise

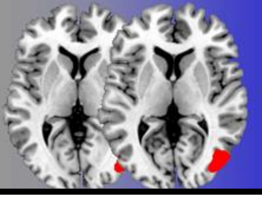
- Noisy brain:
  - Random neural activity
  - signal of interest is relatively weak
- Noisy scanner:
  - Unstructured noise, i.e., measurement error = thermal noise
  - Scanner Drift – the magnetic field can slowly rise and fall
  - Non-uniformities in magnetic field
- Physiological noise:
  - Fluctuations in BOLD due to breathing and heart beat
- Motion
  - head/brain movement due to heartbeat, breathing, subject fidgeting, etc.
- Solutions:
  - Limit subject movement in the scanner
  - **Preprocessing** steps to minimize artifacts and standardize before conducting further analysis

Preprocessing Pipeline



- Preprocessing is performed both on the fMRI data and structural (MRI) scans, collected prior to the experiment.

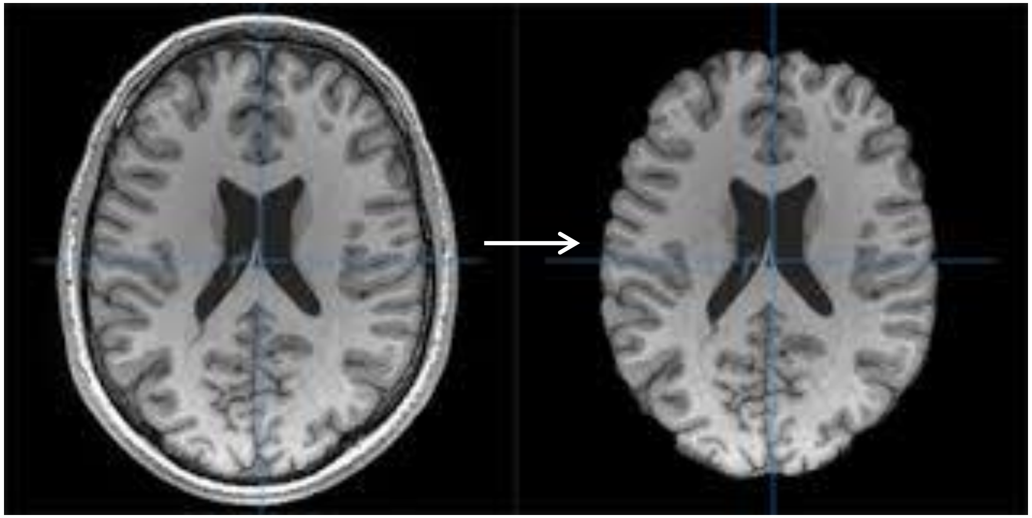
Preprocessing Steps



- Brain Extraction
- Slice timing correction
- Motion correction
- Co-registration
- Normalization
- Spatial Filtering/Smoothing
- Temporal Filtering

Brain Extraction

- Remove non-brain tissue and skull from the image, so that we only use voxels located in the brain.
- Easy to implement with brain extraction tool (BET) in FSL, or 3dSkullStrip in AFNI

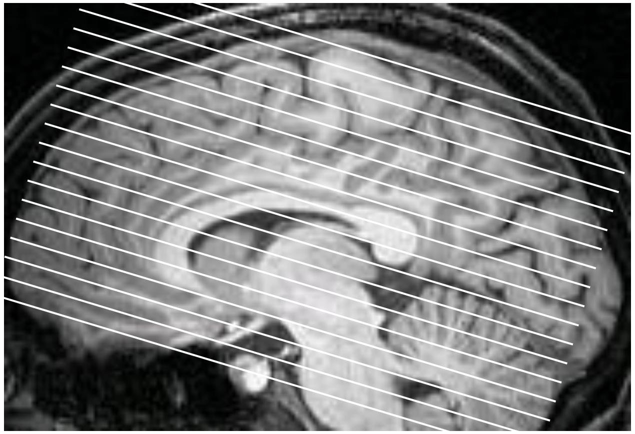


Preprocessing Steps

- Brain Extraction
- Slice timing correction
- Motion correction
- Co-registration
- Normalization
- Spatial Filtering/Smoothing
- Temporal Filtering

Slice Timing Correction

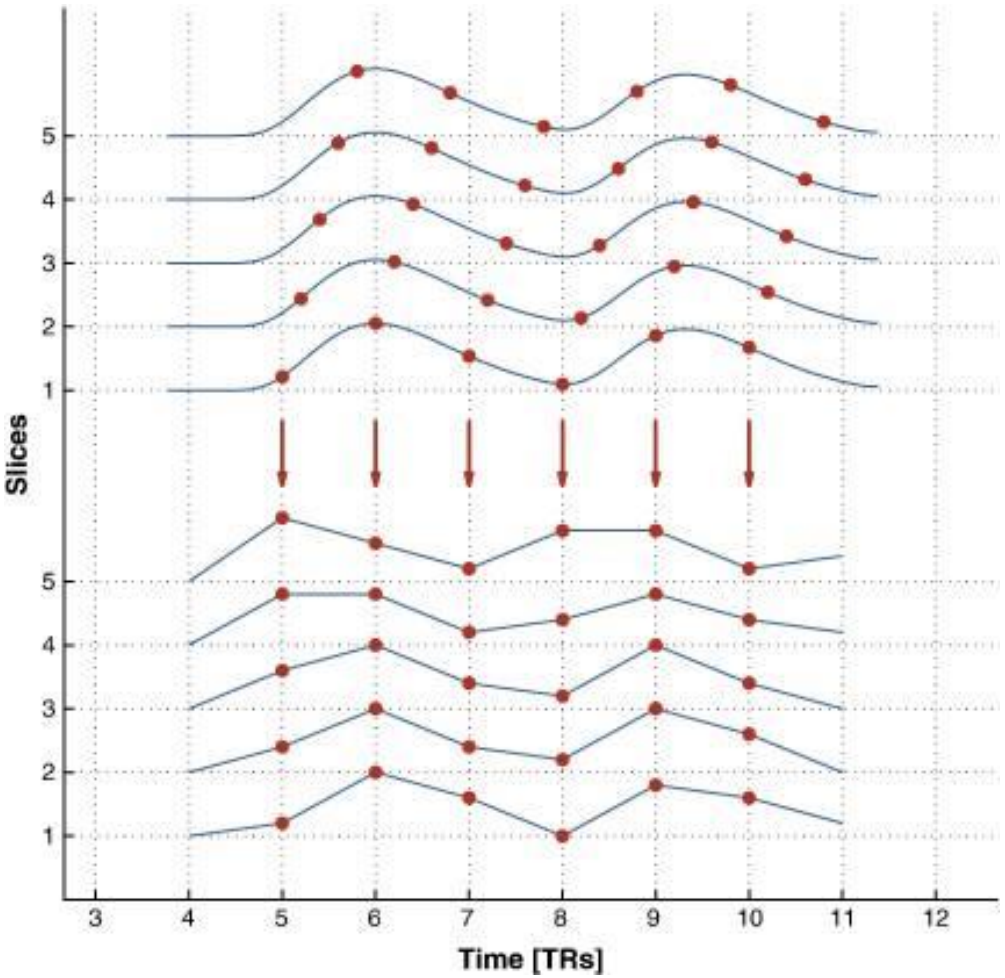
- “Classical” fMRI uses 2D EPI in which a brain volume is acquired in separate slices.
- Each slice is sampled at slightly different time points.
- 2D slices → 3D brain volume



Axial slices

Slice Timing Correction

- **Slice timing correction** shifts each voxel's time series so that they all appear to have been sampled simultaneously.

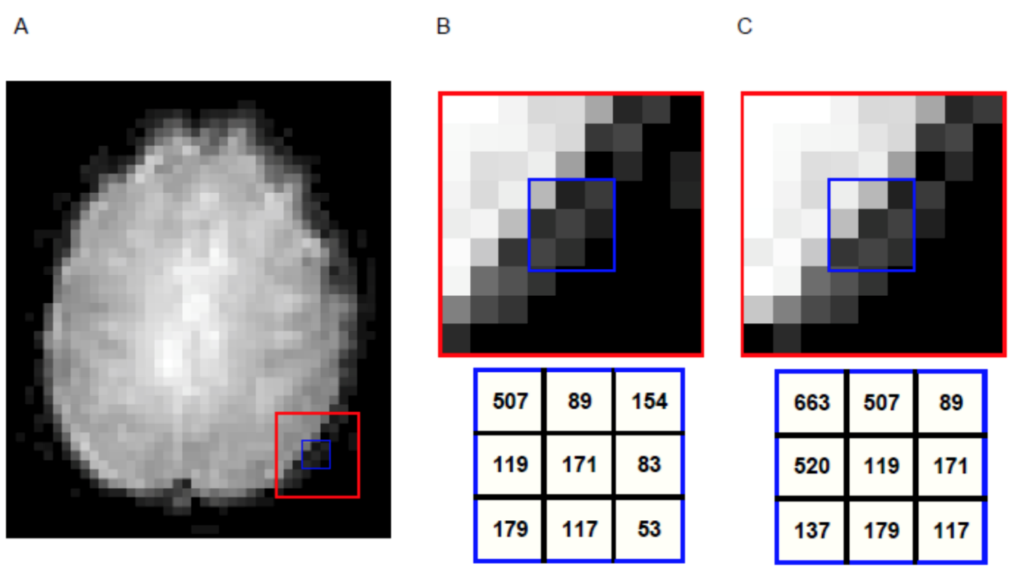


Preprocessing Steps

- Brain Extraction
- Slice timing correction
- Motion correction
- Co-registration
- Normalization
- Spatial Filtering/Smoothing
- Temporal Filtering

Head Motion

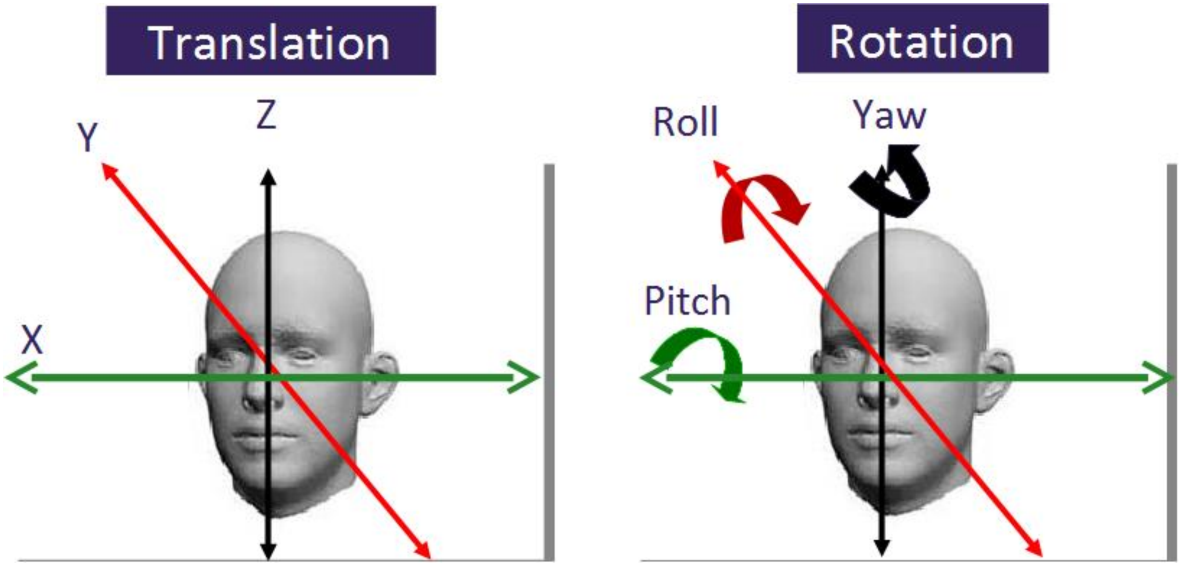
- Small head movements during a scan can be a major source of error if not treated correctly.
- When analyzing a voxel's time series, we assume that the voxel represents the same location in the brain at every time point.
  – Head motion may make this assumption incorrect



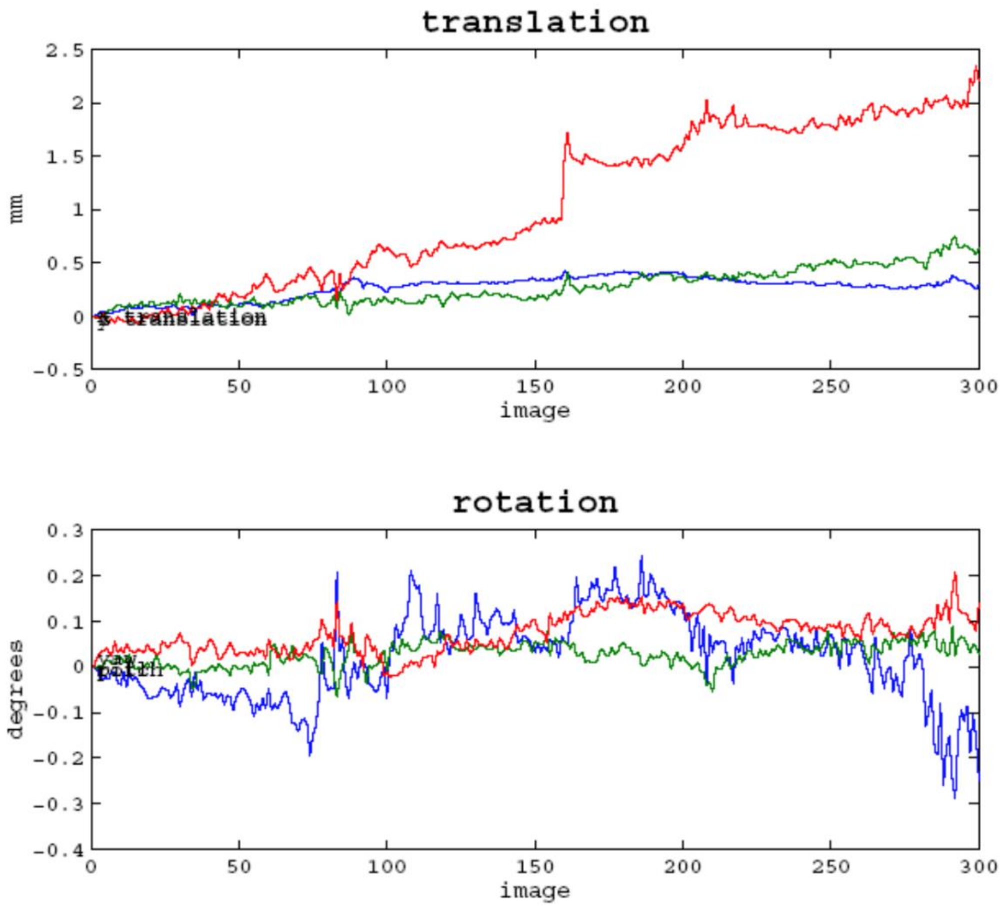
Huettel et al. Functional Magnetic Resonancel Imaging

Motion Correction

- Motion can be corrected using a **rigid body transformation**:
  - Choose a reference volume to register all the other volumes to. (e.g. first volume, middle volume for FSL)
  - Re-aligns to reference volume to minimize variance
  - 6 DOF: translation (x, y, z) and rotation (roll, pitch, yaw)



Motion Correction



Preprocessing Steps

- Brain Extraction
- Slice timing correction
- Motion correction
- Co-registration
- Normalization
- Spatial Filtering/Smoothing
- Temporal Filtering

Coregistration

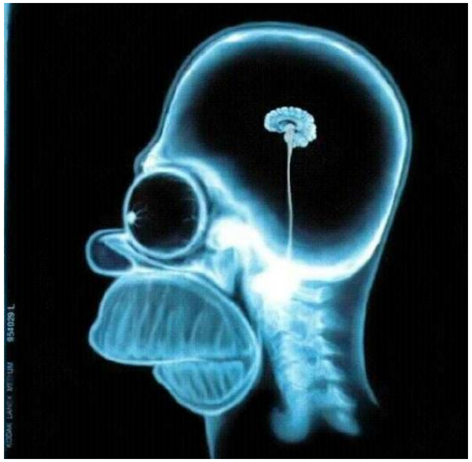
- Functional MRI (T2*) image has low spatial resolution
- It is common to map the results obtained from fMRI onto a high-res structural MRI (T1) image, collected at the start of the scanning session.
- The process of aligning the structural and functional image is called **coregistration**
  - Affine transformation (12 DOF) – translation, rotation, as well as scaling and shearing
  - Non-linear registration allows for warping, can deal with susceptibility artifacts

Preprocessing Steps

- Brain Extraction
- Slice timing correction
- Motion correction
- Co-registration
- Normalization
- Spatial Filtering/Smoothing
- Temporal Filtering

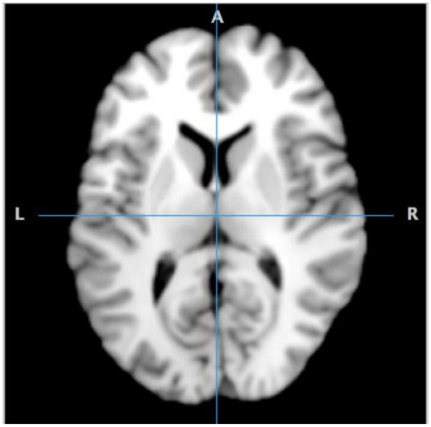
Normalization

- Everyone's brain is different. The brain size of subjects can differ in size by up to 30%!



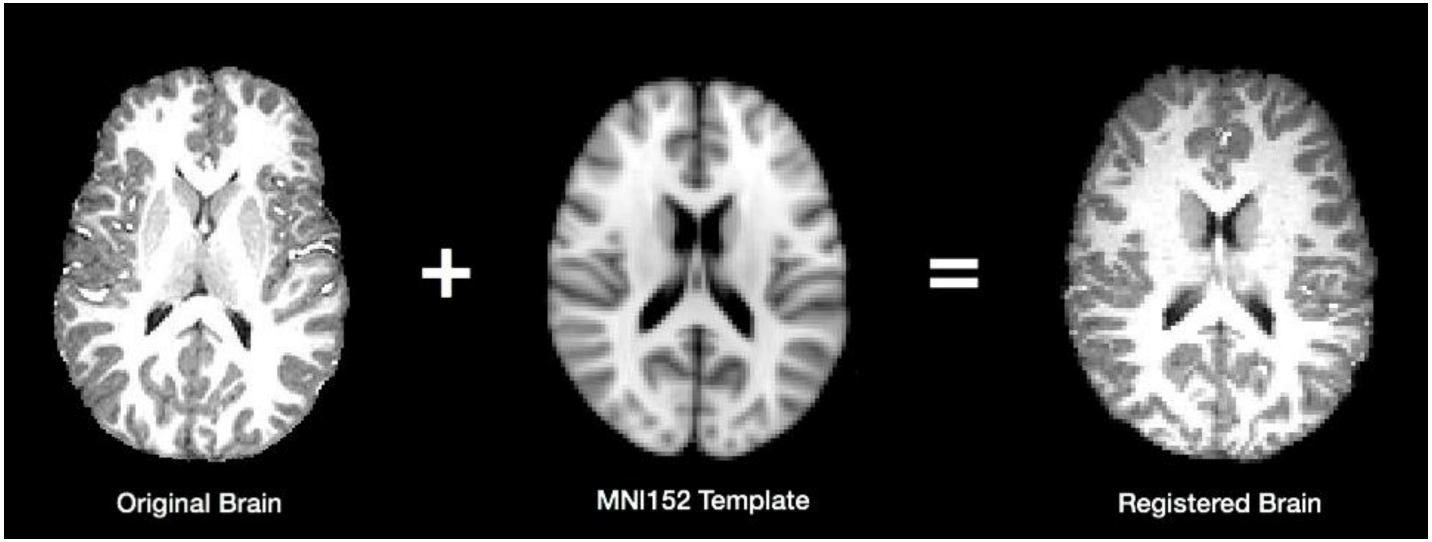
- There is also substantial variation in brain shapes
- **Normalization** attempts to register each subjects anatomy to a standard coordinate space defined by a **template brain**
  - Affine transformation (12 DOF)

Standard Brain Templates



Talairach Template

- **Talairach**
  - Talairach and Tournoux (1988)
  - Based on dissection and photography of a single subject (cadaver of a 60 y.o. female)
- **MNI** (Montreal Neurological Institute)
  - Based on MRI scans of hundreds of normal controls (all RH)



Preprocessing Steps

- Brain Extraction
- Slice timing correction
- Motion correction
- Co-registration
- Normalization
- Spatial Filtering/Smoothing
- Temporal Filtering

Spatial Smoothing

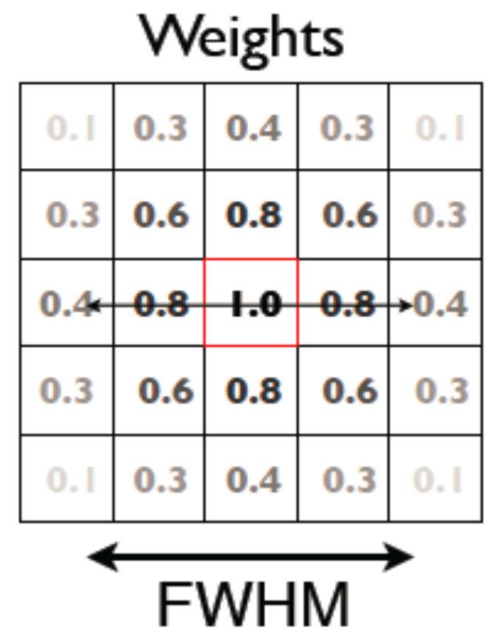
- **Spatial smoothing of fMRI data**: improves inter-subject registration and overcomes limitations in spatial normalization by blurring any residual anatomical differences.

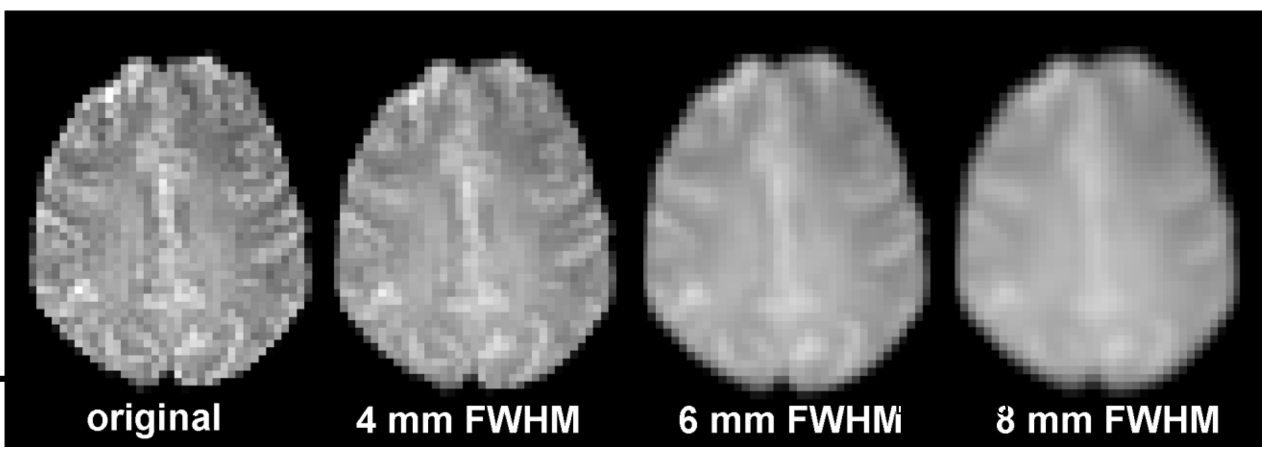
- **PROs**: can increase SNR by decreasing variance and remove artifacts
- **CONs**: may reduce signal if small activations; reduces spatial resolution

- Spatial smoothing is really a bias-variance trade-off: more smoothing = less variance, more bias.

Spatial Smoothing

- Average one voxel's values with its neighbors
- Gaussian Full Width Half Maximum (FWHM) kernel
  - Each voxel intensity is replaced by a weighted average of neighboring intensities
  - Gaussian function specifies weightings and neighborhood size
  - Usu. 4-12 mm FWHM



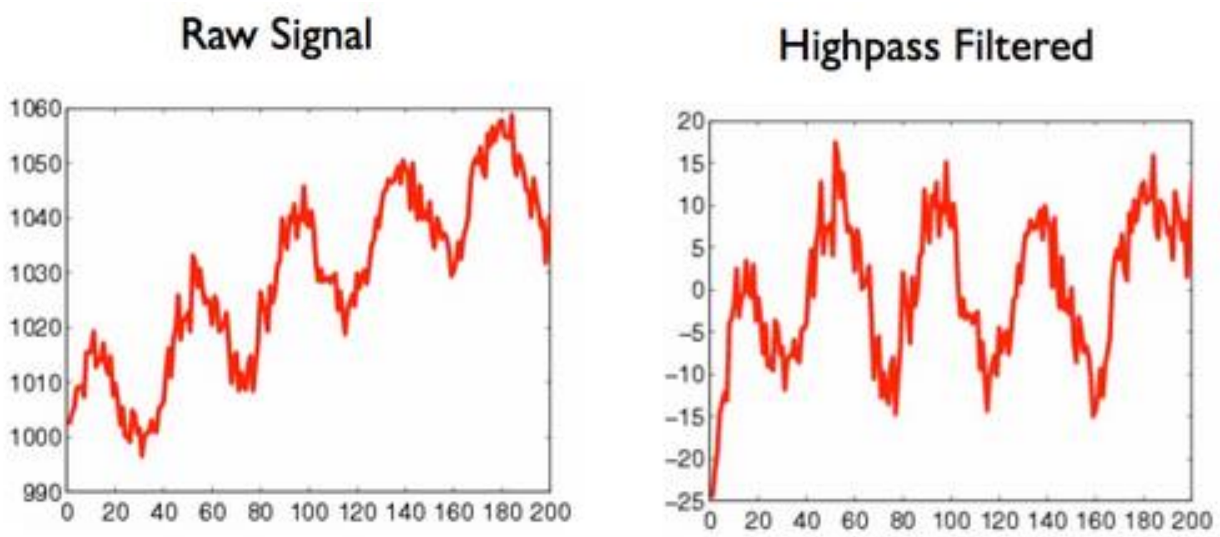


Preprocessing Steps

- Brain Extraction
- Slice timing correction
- Motion correction
- Co-registration
- Normalization
- Spatial Filtering/Smoothing
- Temporal Filtering

Temporal Filtering

- Temporal noise due to drift from scanner, subject's heartbeat and breathing
- Use a high-pass filter to remove low frequency (i.e. long, slow) noise
- Temporal filtering is controversial: introduces correlations

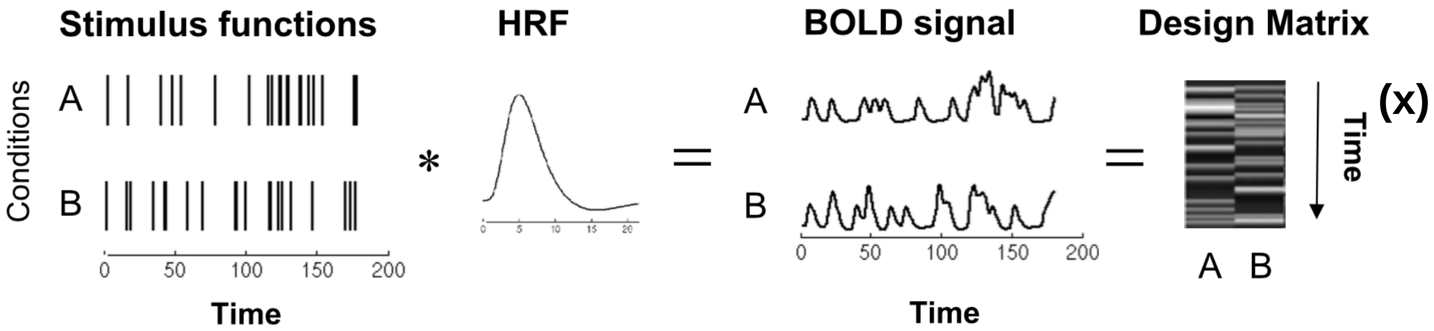


Statistical Analysis

- After the images have been preprocessed, we can begin statistical analysis!
- Goals of statistical analysis of fMRI data:
  - **ACTIVATION**: Localizing brain areas activated by the experimental task
  - **FUNCTIONAL CONNECTIVITY**: Determining networks corresponding to brain function
  - **PREDICTION**: making predictions about psychological or disease states

Activation

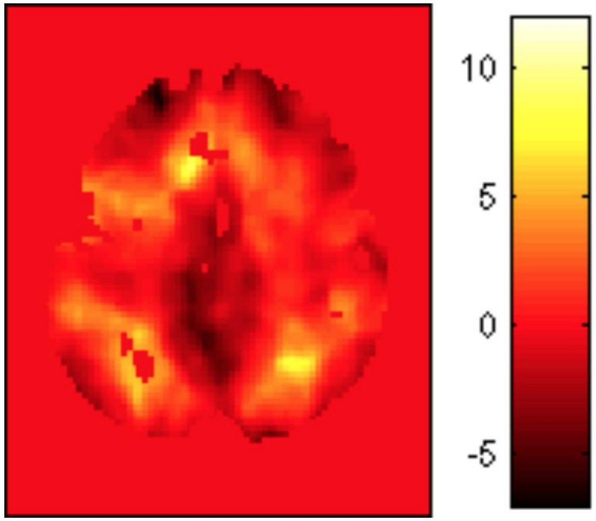
- **Goal**: identify regions that are active during a specific task or related to a certain behavioral measure
- Step 1: construct a model for each voxel
  - "Massive univariate approach"
  - Regression models (General LM) commonly used



$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \qquad \boldsymbol{\varepsilon} \sim N(\mathbf{0}, \mathbf{V})$$

- Step 2: perform a statistical test to determine whether task-related activation is present in each voxel

$$H_0 : \mathbf{c}^T \boldsymbol{\beta} = 0$$

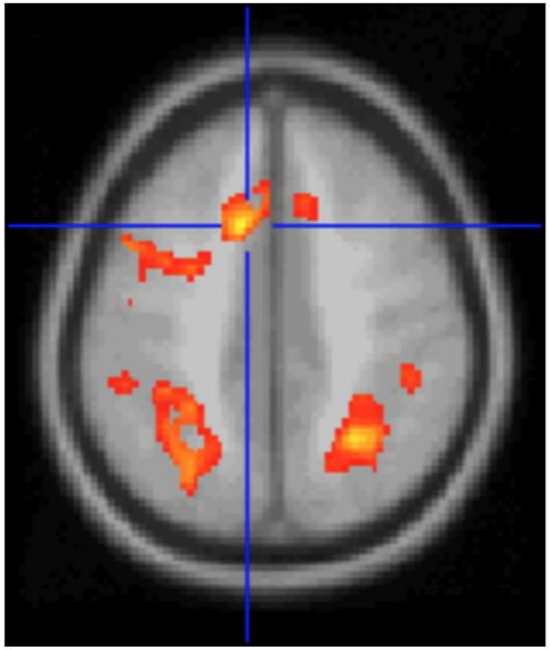


For contrast c, task vs. control

Statistical map: map of t-test statistics across all voxels (a.k.a. t-map)

Activation

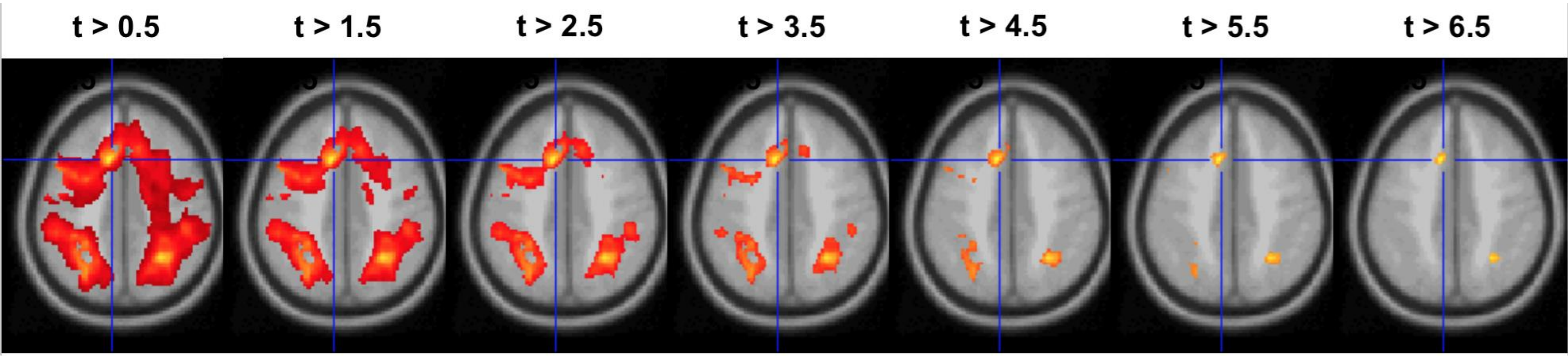
- Step 3: Choose an appropriate threshold for determining statistical significance



Thresholded t-map: Each significant voxel is color-coded according to the size of its p-value

Multiple Comparison Problem

- Which of 100,000 voxels are significant?
  - α=0.05 → 5,000 false positive voxels
- Bonferroni correction is overly conservative
- Choosing a threshold is a balance between sensitivity (true positive rate) and specificity (true negative rate)



Methods for thresholding

- The Bonferroni correction
- Random Field Theory
- Permutation Tests
- False Discovery Rate (FDR)

Family-Wise Error (FWE)

given that the whole family of the test statistics is from the null distribution, the probability of there being one or more test statistic values that exceed a pre-specified threshold.

B. Risk, D. Matteson, N. Spreng, D. Ruppert

SPATIOTEMPORAL MIXED MODEL

EMORY ROLLINS SCHOOL OF PUBLIC HEALTH | CBIS Center for Biomedical Imaging Statistics

Our approach: STMM [Risk et al., 2016]

- We model each region independently: $r \in \{1, \ldots, R\}$ index region using [Gordon et al., 2016] (172 parcels / hemisphere)
- $\mathbf{Y}_{ir} = [y_{ir11}, \ldots, y_{irV_rT}]'$

$$\mathbf{Y}_{ir} = \mathbf{1}_{V_r} \otimes \mathbf{X}_i \boldsymbol{\beta}_{r\cdot} + (\mathbf{I}_{V_r} \otimes \mathbf{X}_i)\, \boldsymbol{\beta}_r + \mathbf{1}_{V_r} \otimes \mathbf{X}_i \boldsymbol{s}_{ir} + (\mathbf{I}_{V_r} \otimes \mathbf{X}_i)\, \boldsymbol{b}_{ir} + (\mathbf{I}_{V_r} \otimes \mathbf{Z}_i)\, \boldsymbol{\gamma}_{ir} + \boldsymbol{e}_{ir}$$

- $\boldsymbol{s}_{ir} \overset{iid}{\sim} \mathcal{N}(\mathbf{0}_Q, \mathbf{S}_r)$ with $\mathbf{S}_r = \mathrm{diag}(\sigma^2_{s_{r1}}, \ldots, \sigma^2_{s_{rQ}})$
- $\boldsymbol{b}^q_{ir} \overset{iid}{\sim} \mathcal{N}(\mathbf{0}_{V_r}, \sigma^2_{b_{rq}} \Omega_{rq})$
- $\boldsymbol{e}_{ir} \sim \mathcal{N}(\mathbf{0}_{V_rT}, \oplus_{v=1}^{V_r} \xi^2_{irv} \boldsymbol{\Psi}_{irv})$ where $\oplus_{v=1}^{V_r} \xi^2_{irv} \boldsymbol{\Psi}_{irv} = \mathrm{diag}(\xi^2_{ir1} \boldsymbol{\Psi}_{ir1}, \ldots, \xi^2_{irV_r} \boldsymbol{\Psi}_{irV_r})$
- Subject activation is defined

$$a_{irvq} = \beta_{r\cdot q} + \beta_{rvq} + s_{irq} + b_{irvq}$$

STMM

Visualizing STMM

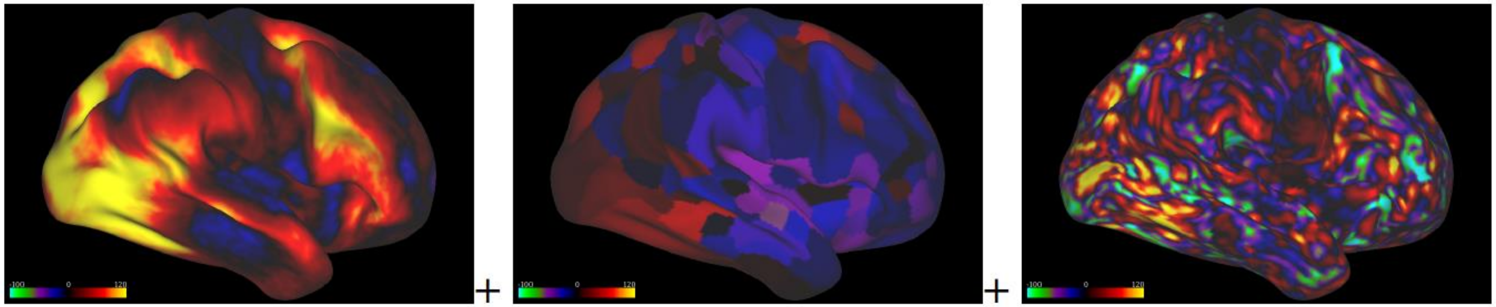


Figure : Population regional + vertex effects, $\beta_{r\cdot q} + \beta_{rvq}$ (left)
Subject regional, s_{irq} (middle)
Subject-vertex, b_{irvq} (right)

=

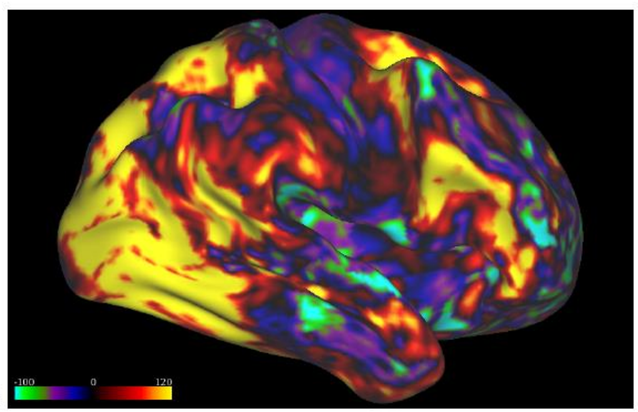


Figure : Subject activation: a_{irvq}

HCP Analysis

MUMM subject activation for HCP Motor Task

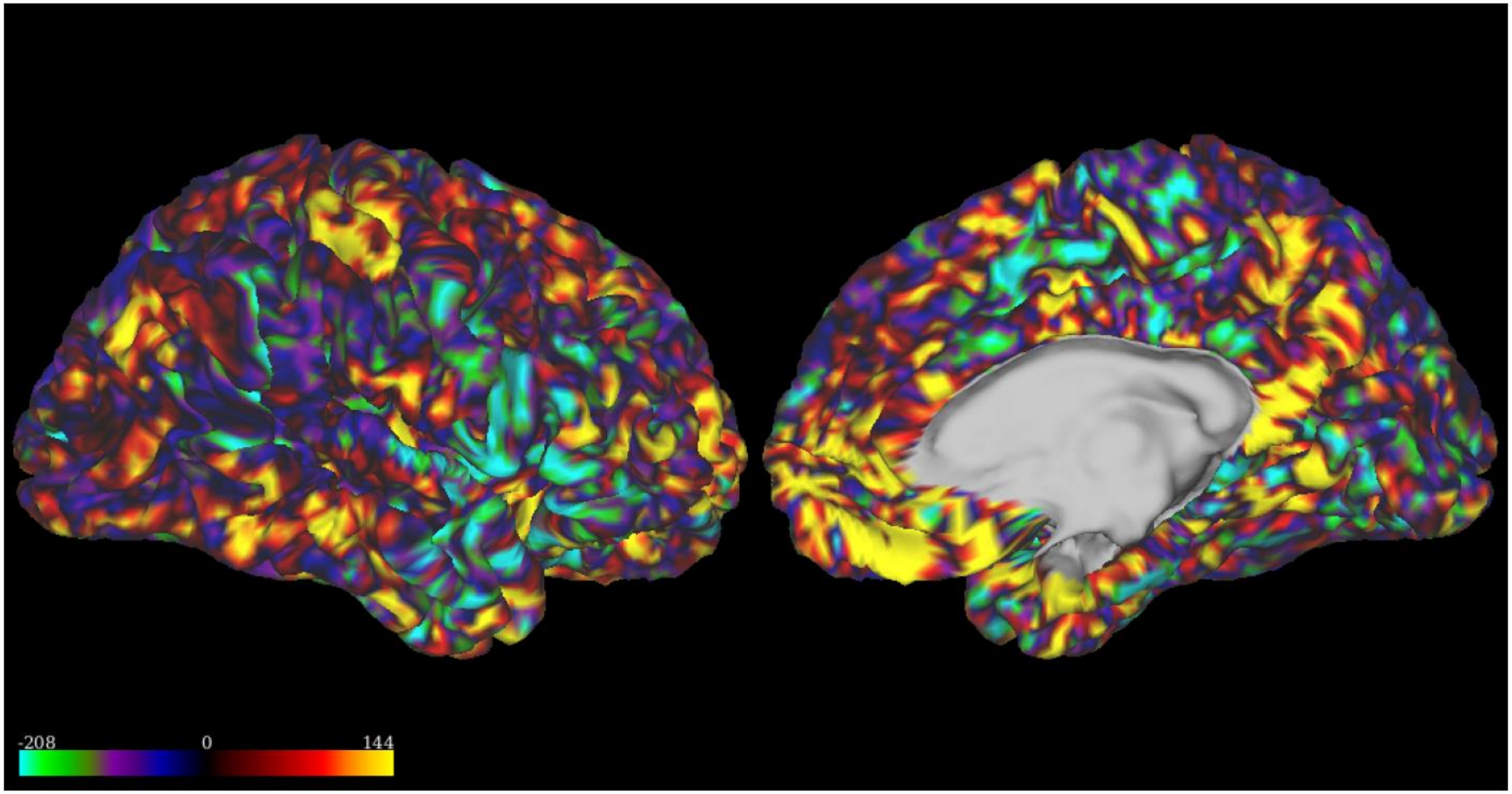


Figure : MUMM estimates of the contrast between the left-hand finger tap vs other tasks for a randomly selected subject (123925) (right hemisphere).

HCP Analysis

STMM subject activation for HCP Motor Task

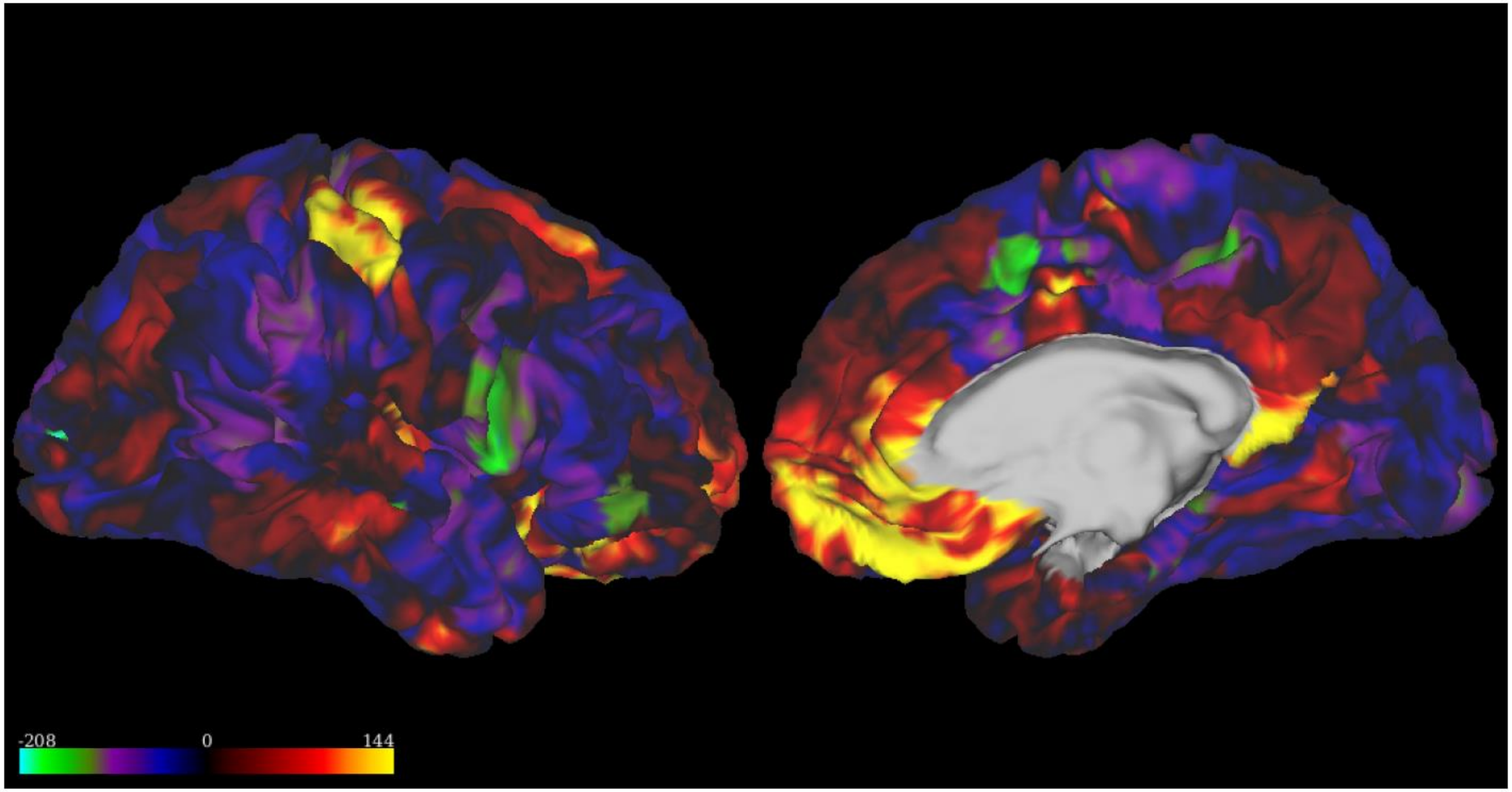


Figure : STMM estimates of the contrast between the left-hand finger tap vs other tasks for a randomly selected subject (123925) (right hemisphere).

Statistical Analysis

- Goals of statistical analysis of fMRI data:

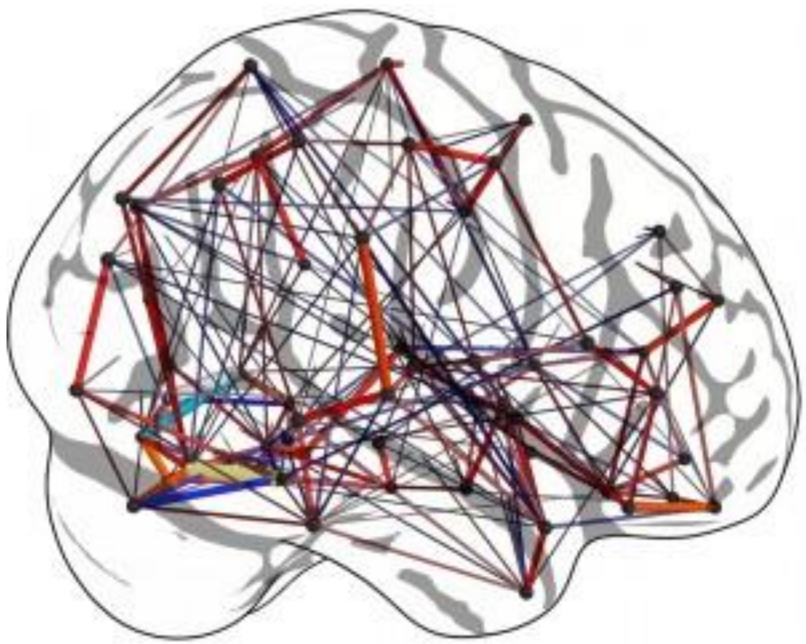
  – **ACTIVATION**: Localizing brain areas activated by the experimental task

  – **FUNCTIONAL CONNECTIVITY**: Determining networks corresponding to brain function

  – **PREDICTION**: making predictions about psychological or disease states

Connectivity

- Recently, there has been increased interest in augmenting activation analyses with **connectivity studies**, which describe how various brain regions interact.

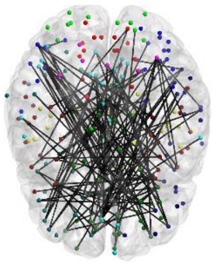


Connectivity

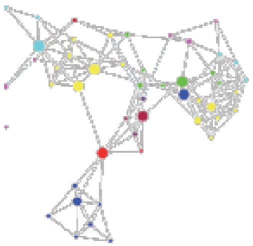
BRAIN CONNECTIVITY

•**Structural connectivity**
  - Diffusion MRI tractography

•**Functional connectivity**
  - seed-based analysis, graphical models, ICA



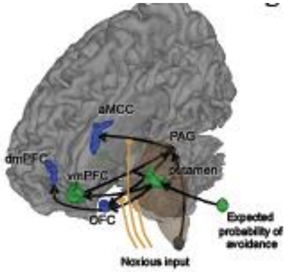
Wang et al. 2016



Wager et al. 2015 graphical model

•**Effective connectivity**
  -Granger causality, Dynamic Causal Modeling (DCM)

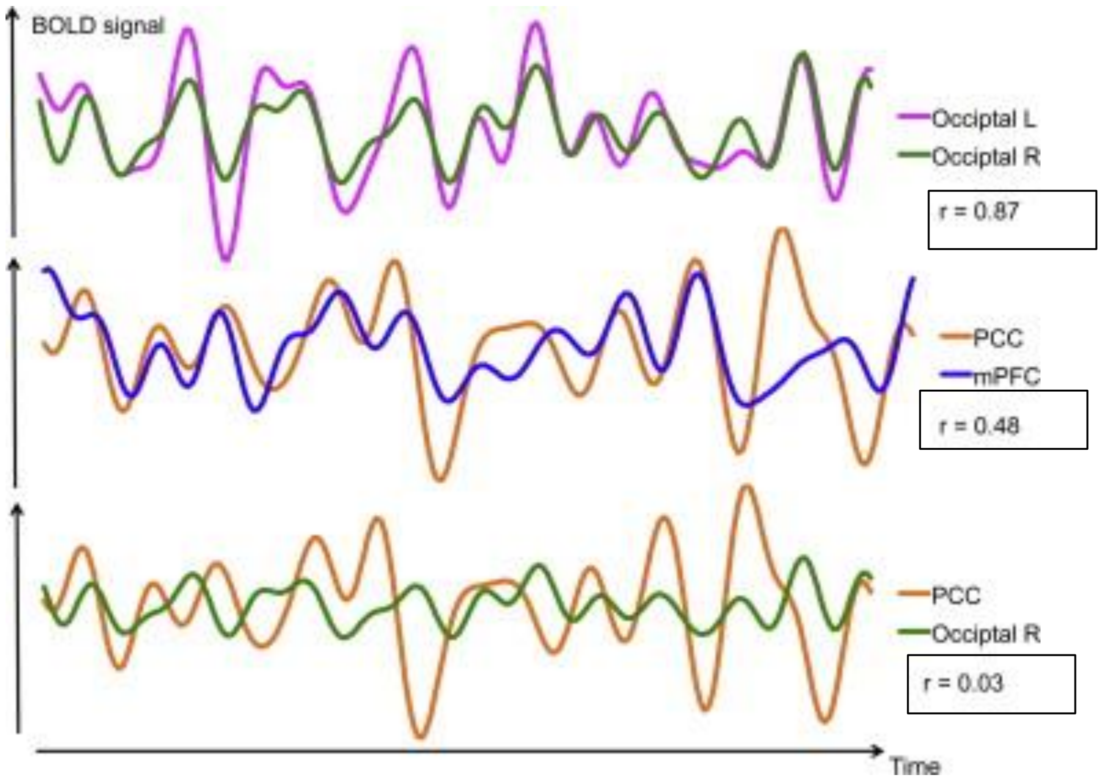
•**Dynamic connectivity**
  -sliding window, hidden Markov model, change-point method



Roy et al. 2014 DCM

Functional Connectivity

- Temporal coherence of brain regions
- Undirected association (usu. correlation)



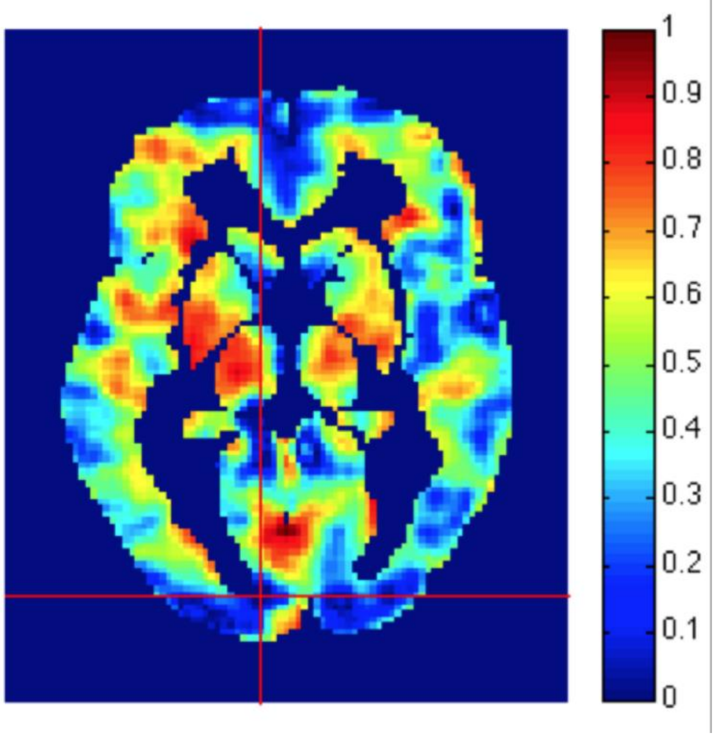
Functional Connectivity

- **Functional Connectivity Analysis** is usually performed using data-driven methods which make no assumptions about the underlying biology

- Methods include:
  - Seed analysis
  - Network analysis
  - Partitioning methods: Clustering, PCA, ICA

Seed Analysis

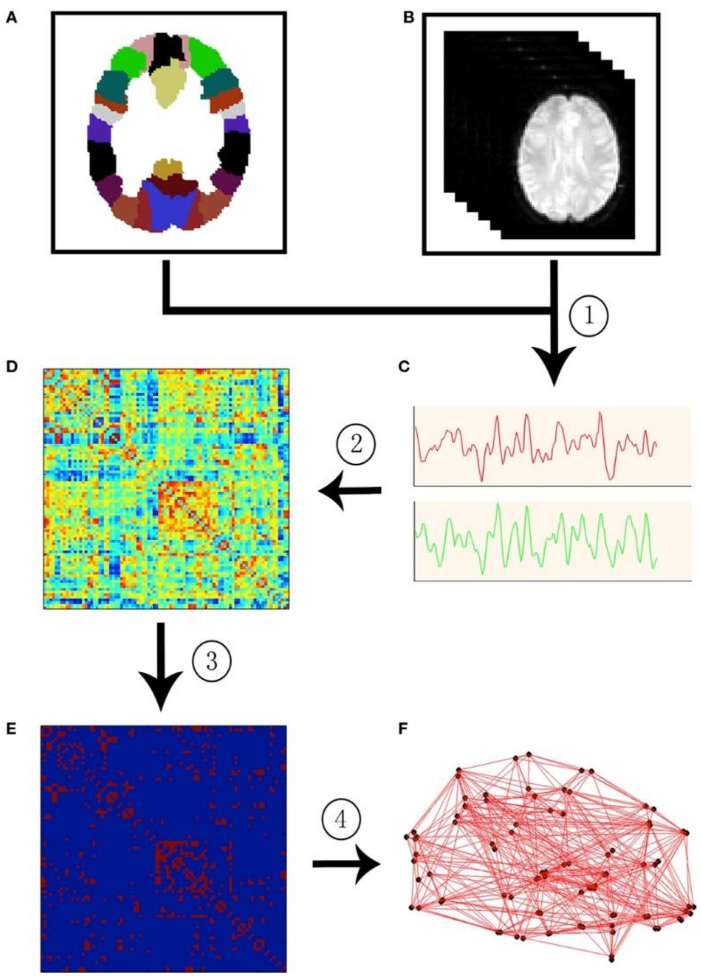
- Calculates the correlation between the temporal brain activity profile in a selected (“seed”) voxel/region and the profiles from other voxels/regions in the brain.
- Simple and easy to implement
- BUT… requires careful selection of seed voxel/region
- Provides a limited view of the brain, since it is restricted to connectivity involving the seed voxel.



Network Analysis

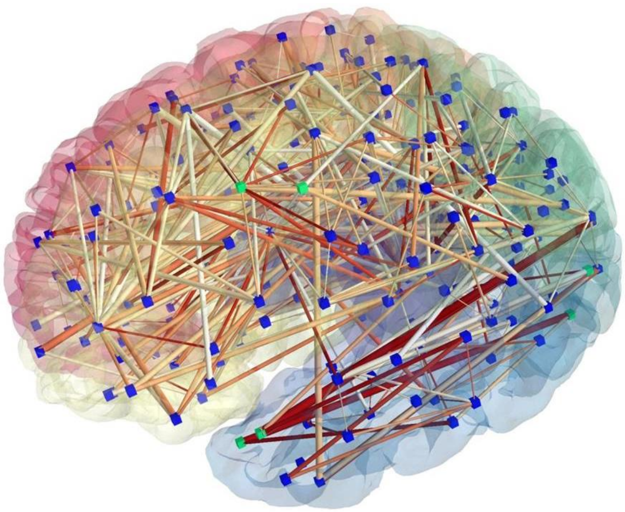
Whole-Brain region-to-region approach:

1. Parcellate the brain and extract the “average” fMRI time course for each region
2. Calculate correlation between regions → correlation matrix
3. Threshold → binary adjacency matrix
4. Graph theory analysis



Network Analysis

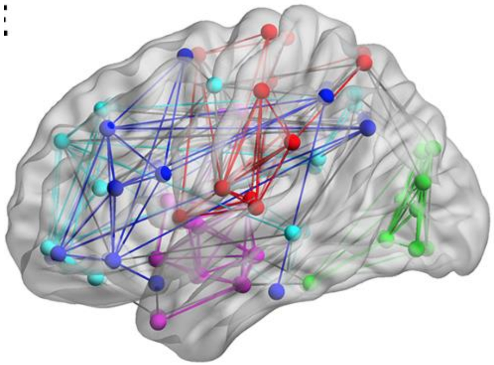
- **Network/Graph Theory analysis** tries to characterize networks using a small number of meaningful summary measures
- Comparing network topological measures (ex: node degree, clustering coefficient, etc.) between groups of subjects may reveal connectivity abnormalities related to brain disorders



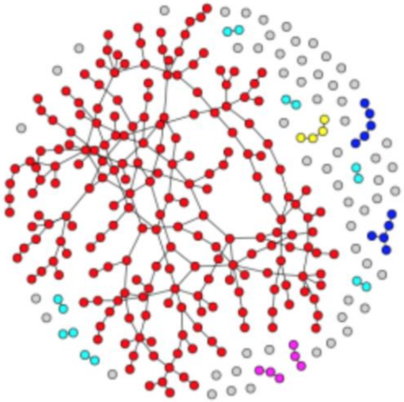
A network is a system of **nodes** (regions) and **edges** (connections between regions)

Network Analysis

- Network Visualization tools
  - BrainNet MATLAB toolbox



  - igraph R package



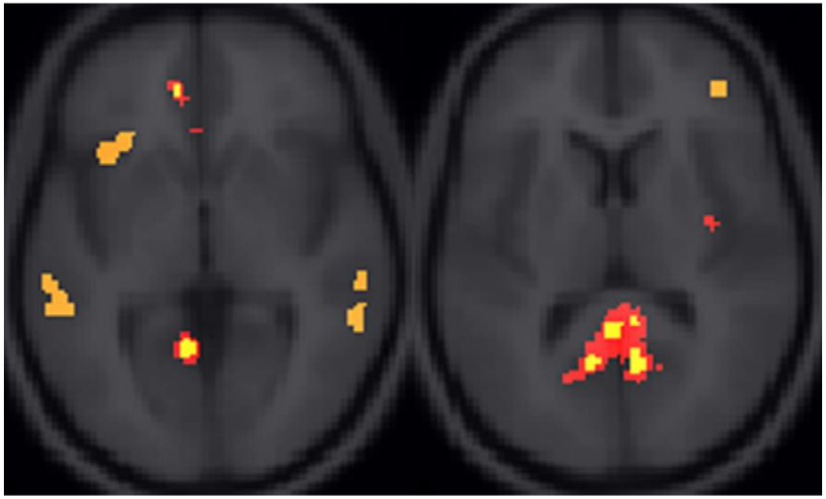
- Brain connectivity toolbox (MATLAB) for calculating graph theory metrics to characterize networks.

Partitioning Algorithms

- **Partitioning algorithms** identify spatially distinct components or clusters in the brain
- Each of these components represents a functionally connected network
- Methods:
  - Clustering
  - PCA
  - ICA

Clustering

- **Cluster analysis:** identifies "clusters" of voxels with similar brain activity patterns.
- Clusters may consist of noncontiguous voxels, offering the potential of identifying associations between anatomically distant voxels
- Several algorithms: K-means approach, fuzzy clustering, hierarchical clustering, etc.



3 clusters (red, orange, yellow) based on mean brain activity of cocaine addicts in inhibitory control study.

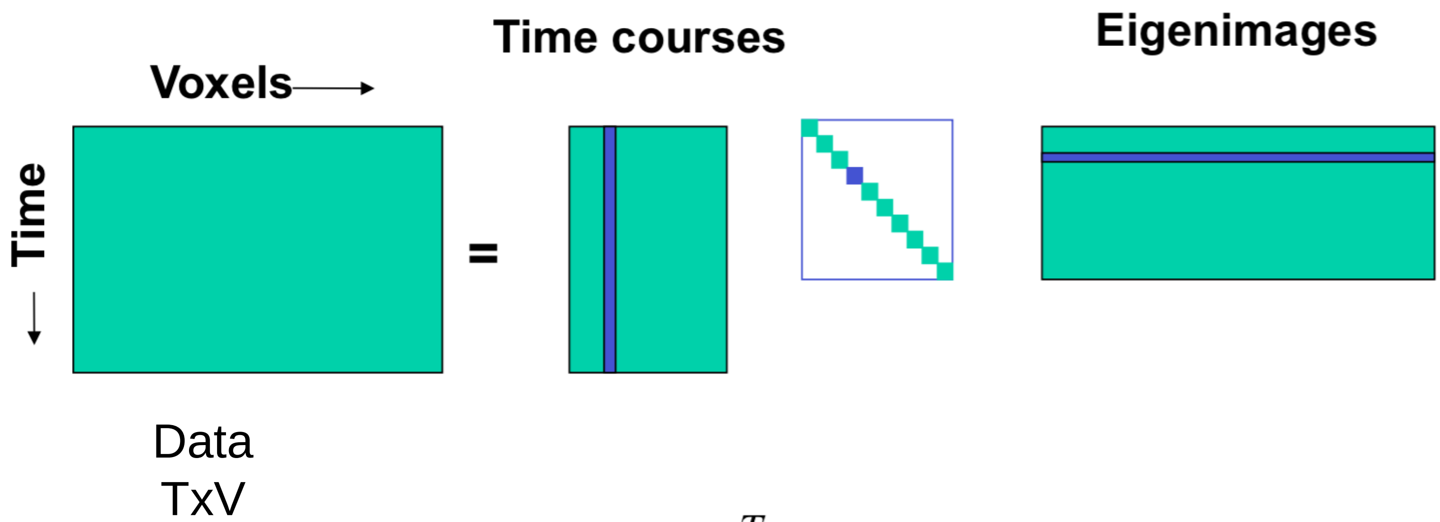
Each cluster contains voxels with similar patterns of brain activity

PCA

- **Principal components analysis (PCA)** involves finding spatial modes, or eigenimages, in the data
  - These are the patterns that account for most of the variance-covariance structure in the data, ranked in order

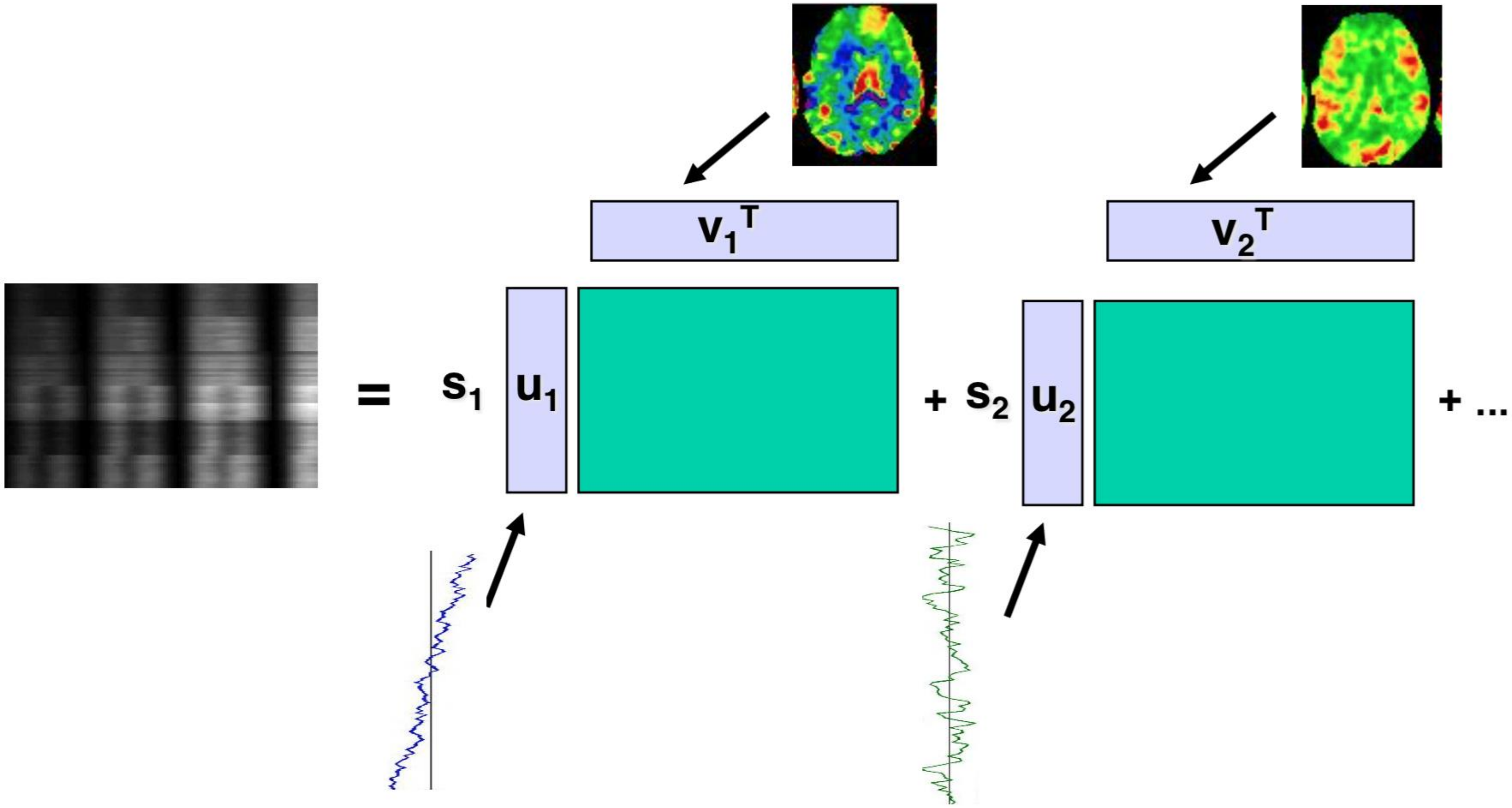
- The eigenimages can be obtained using singular value decomposition (SVD), which decomposes the data into two sets of orthogonal vectors that correspond to patterns in space and time.

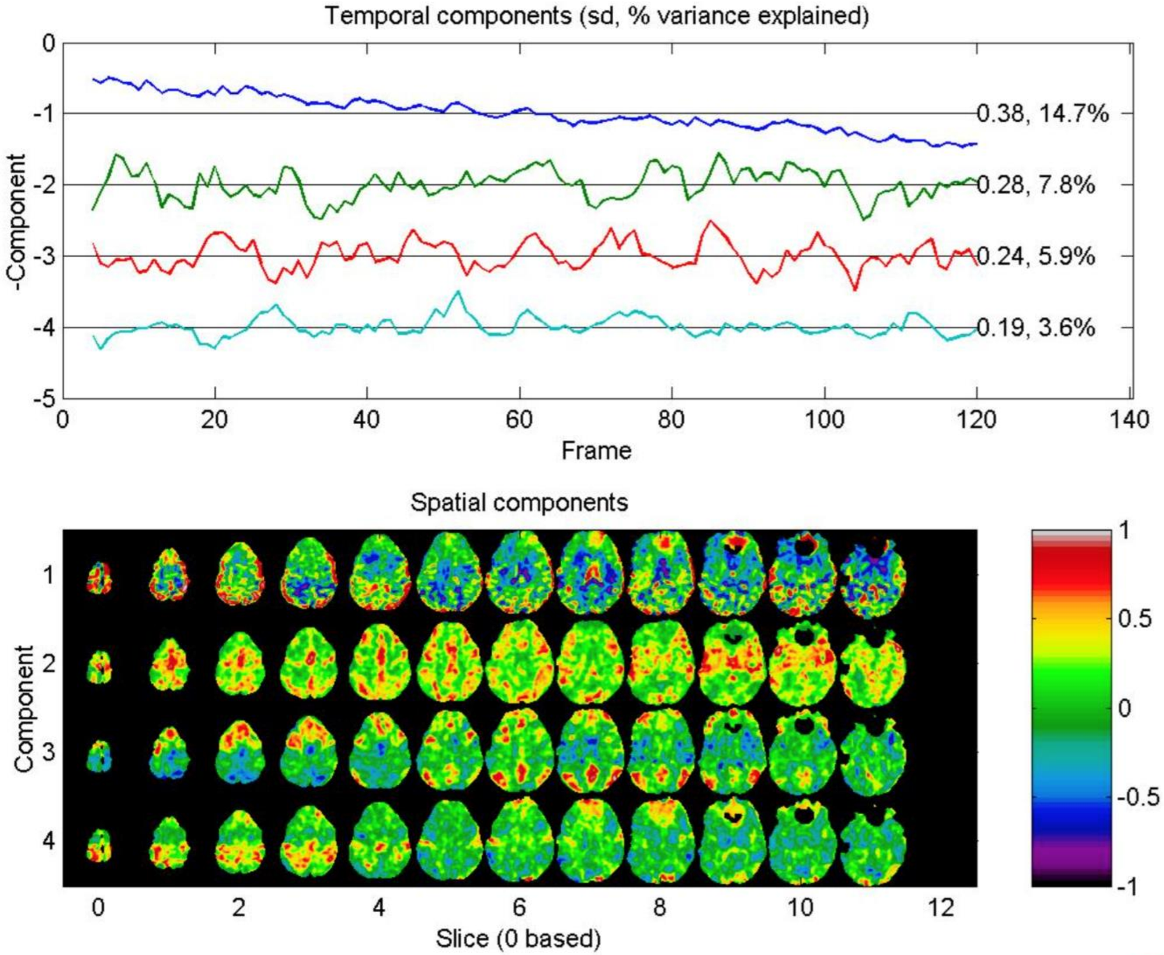
PCA



$$\mathbf{X} = \mathbf{USV}^T$$

$$\mathbf{X} = s_1 \mathbf{u}_1 \mathbf{v}_1^T + s_2 \mathbf{u}_2 \mathbf{v}_2^T + \ldots + s_N \mathbf{u}_N \mathbf{v}_N^T$$


=
s_1
u_1
v_1^T
+ s_2
u_2
v_2^T
+ ...


Temporal components (sd, % variance explained)
-Component
Frame
0.38, 14.7%
0.28, 7.8%
0.24, 5.9%
0.19, 3.6%
0
-1
-2
-3
-4
-5
0
20
40
60
80
100
120
140
Spatial components
Component
Slice (0 based)
1
2
3
4
0
2
4
6
8
10
12
1
0.5
0
-0.5
-1

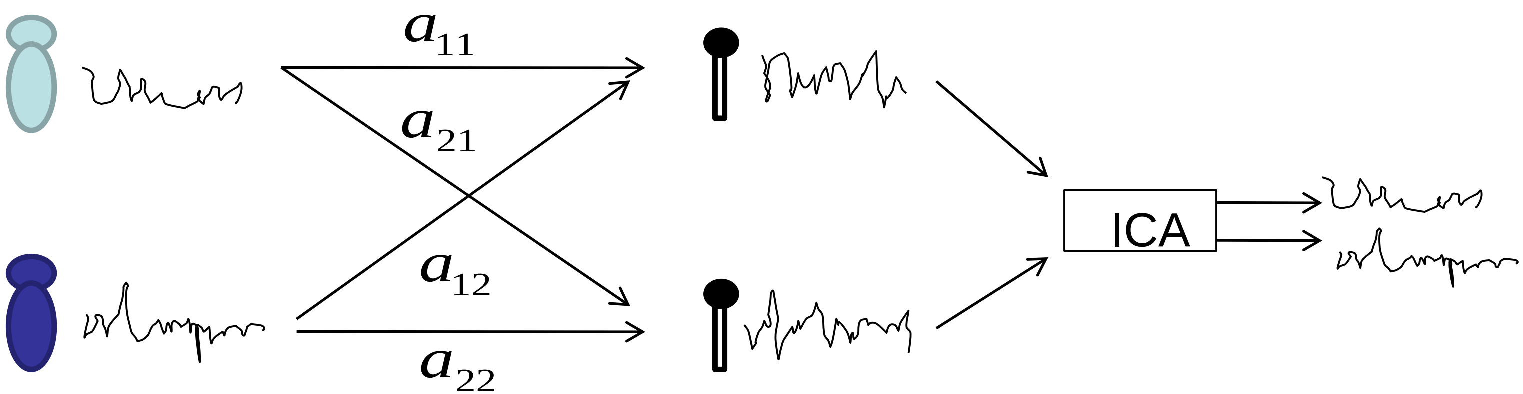
Worsley

ICA

- Definition of ICA

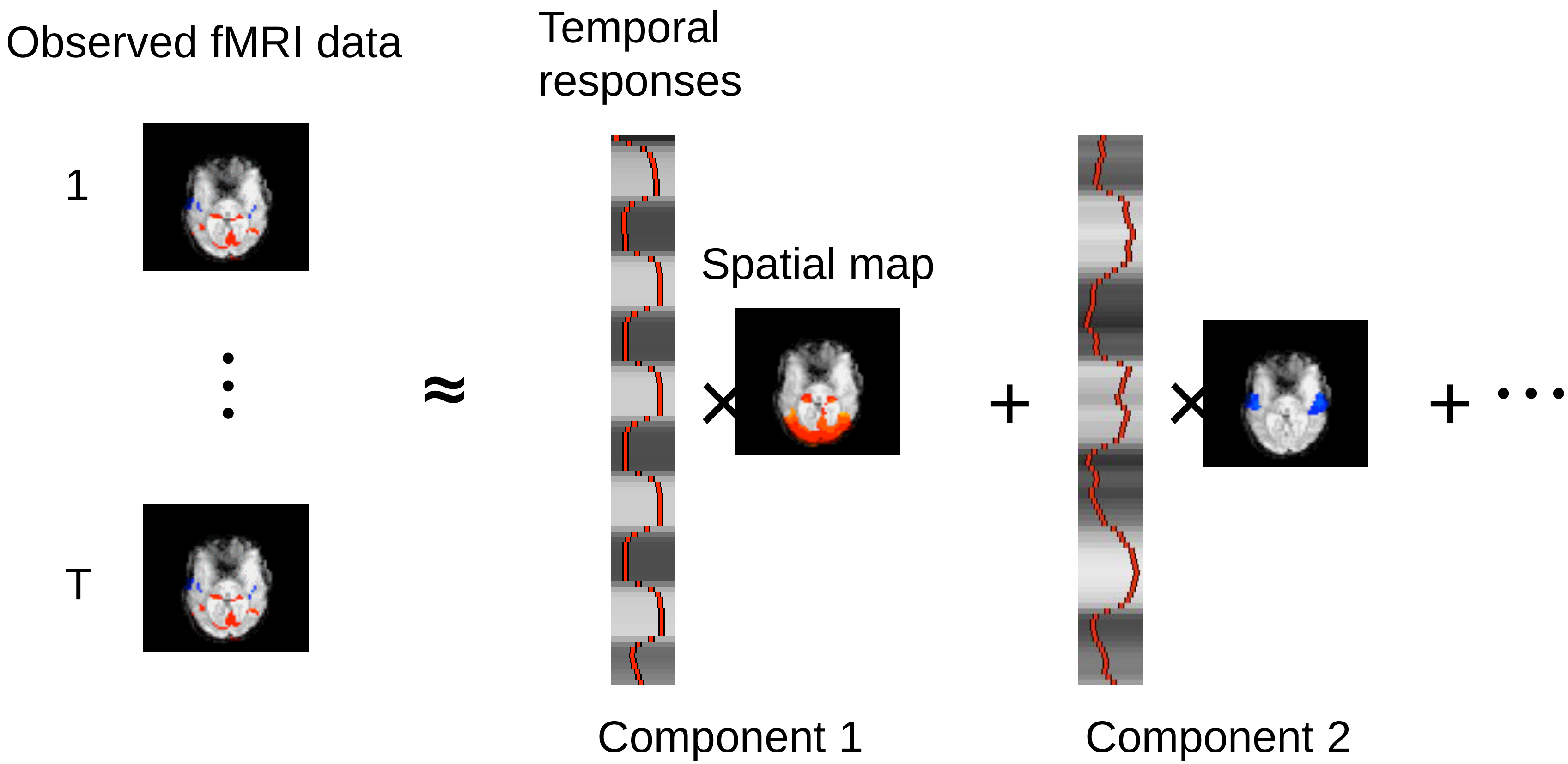
Independent component analysis (**ICA**) is a computational method for separating a multivariate signal into additive subcomponents assuming the mutual statistical independence of the non-Gaussian source signals. It is a special case of blind source separation.

- Classic Example: Cocktail Party Problem



- Key assumptions of ICA
  - signals are statistically independent
  - signals are non-Gaussian
  - # of mixture of signals ≥ # of sources

ICA

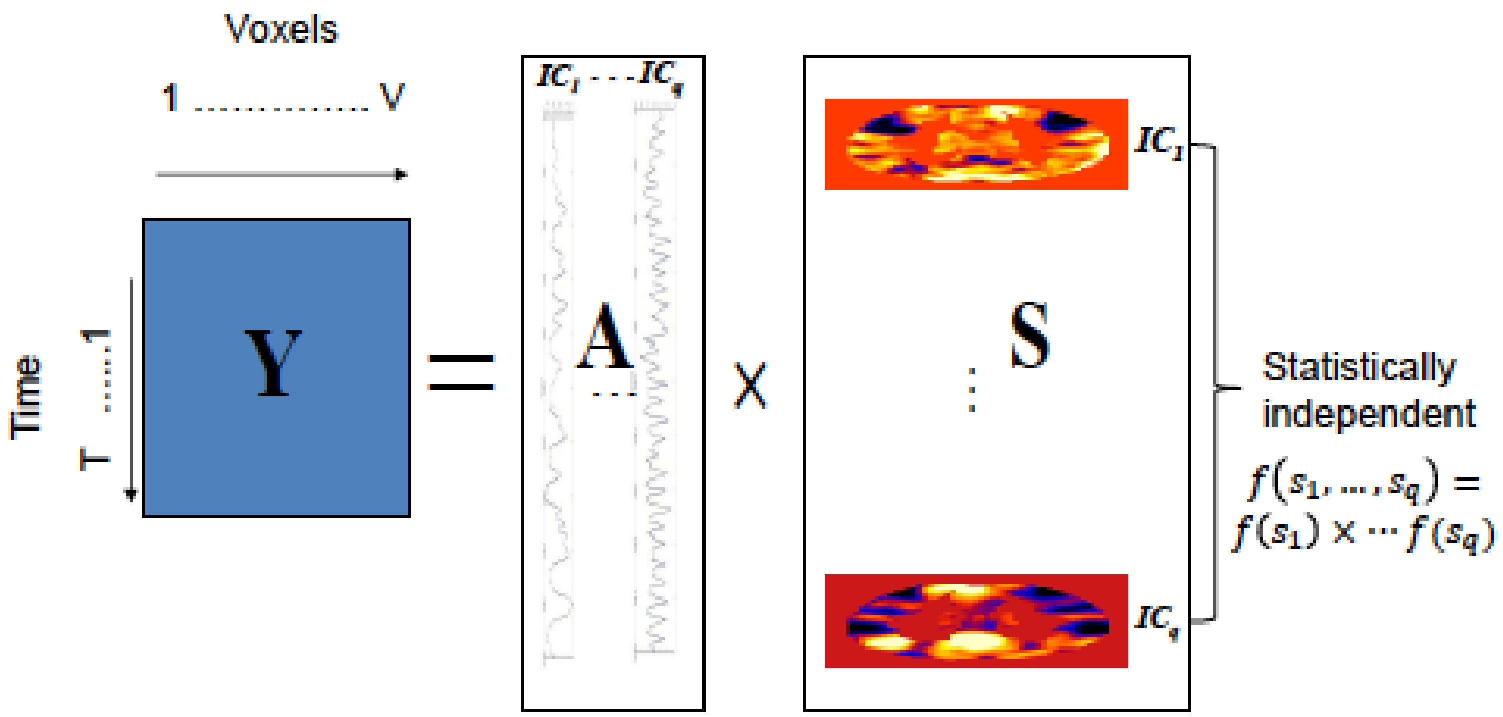


Goal: Decompose observed fMRI data as a linear combination of spatio-temporal processes of underlying source signals.

Figure: MELODIC at http://www.fmrib.ox.ac.uk/analysis/research/melodic/

Standard ICA for fMRI

$$Y_{T\times J} = \mathbf{A}_{T\times q}\,\mathbf{S}_{q\times J}$$



Advantages of ICA

- Does not require any a priori assumptions about the spatiotemporal structure underlying the observed brain activity
- Can be used for fMRI data with any paradigm; esp. useful for resting-state data where no clear task-related activations exist
- Simultaneously separates neuronal and non-neuronal sources (e.g. respiration) into different components
- ICA is more effective than PCA at identifying functional networks (Beckmann et al, 2005). Uses high-order statistics from the data
- Easy to extend to multi-subject case for group inference – GIFT, FSL Melodic toolboxes, HINT (by CBIS, upcoming)

A partial correlation method for whole brain network modeling

DensParcorr developed by CBIS (Wang et al., 2016, available from CRANS)

Dens-based Partial Correlation Estimation Approach

Step 1. Initial: Calculate the sample covariance matrix $\widehat{\boldsymbol{\Sigma}}$ based on the observed fMRI time series from M nodes in the brain. If one would like to impose sparsity regularization on the precision matrix estimate, specify a percentage p, where $p \in (0,1)$, for selecting the tuning parameter based on the desired dense level of the precision matrix estimate.

Step 2. choose sparsity tuning parameter

$$\lambda_{\mathrm{p}}^{*} = \operatorname{argmin}_{\lambda_{\mathrm{n}}} \{|Dens(\lambda_{\mathrm{n}}) - p \times Dens_{\max}|\}$$

Step 3: Estimate the precision matrix using CLIME $\Omega^{*}(\lambda)$

Step 4: Derive estimate for the partial correlation matrix

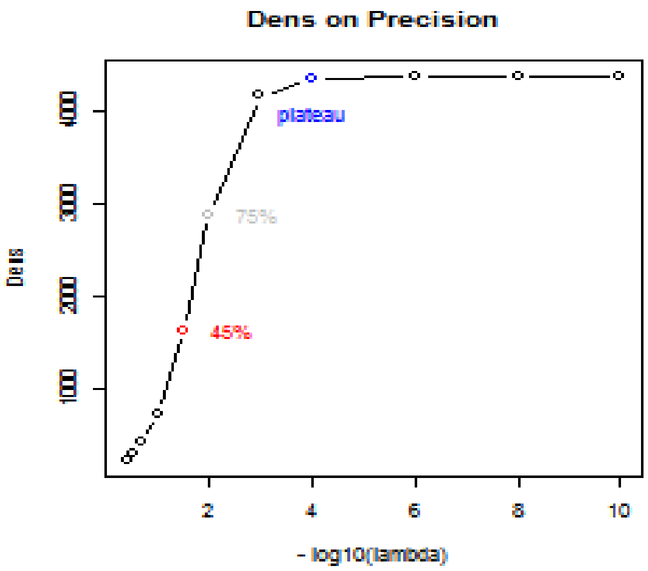
$$\mathbf{Pcorr} = -\mathrm{diag}(\boldsymbol{\Omega}^{*})^{-1/2}\boldsymbol{\Omega}^{*}\mathrm{diag}(\boldsymbol{\Omega}^{*})^{-1/2} + 2\mathbf{I}_{\mathrm{M}}$$

Dens-based tuning parameter selection method

To measure how dense an estimated precision matrix is, we propose the following *Dens* criterion function,

$$Dens(\boldsymbol{\Omega}) = \sum_{ij} |\omega_{ij}| \text{ , where } \boldsymbol{\Omega} = \{\omega_{ij}\}$$

Essentially, *Dens* is the matrix-wise L1 norm of $\boldsymbol{\Omega}$.



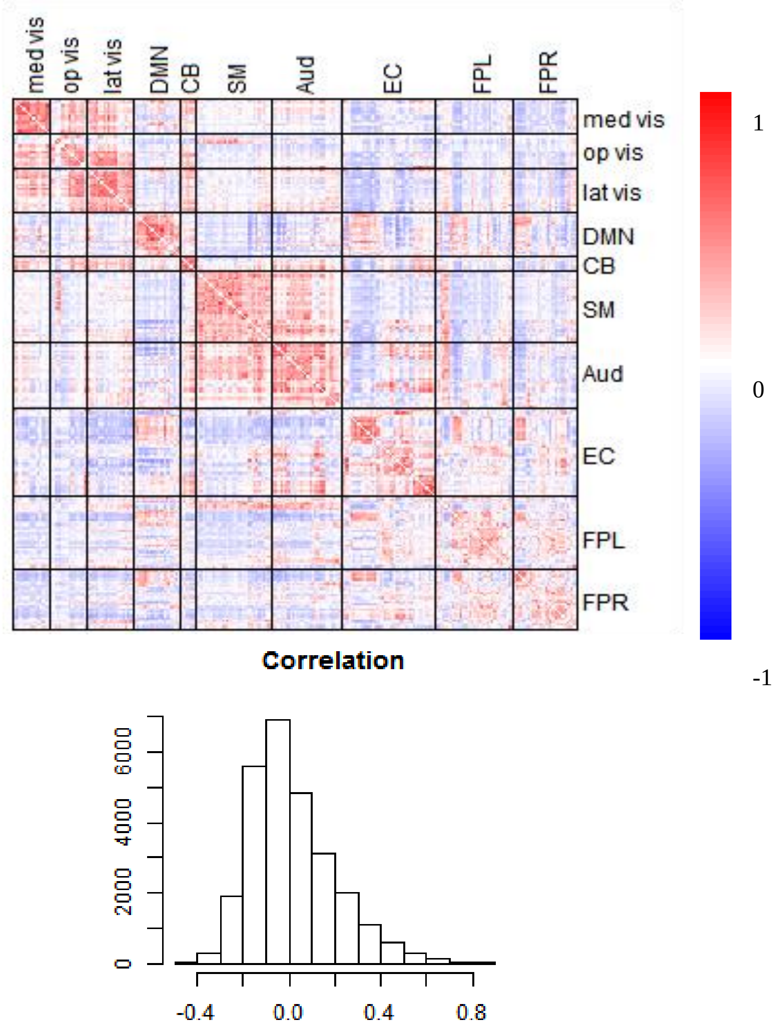
- Specify a monotonically decreasing sequence $\{\lambda_n, n = 0,1,\ldots,\}$ within the range (0,1) with $\lambda_0 \to 1$ and $\lambda_n \to 0$ as n increases
- Obtain CLIME $\boldsymbol{\Omega}$ estimate for $\{\lambda_n\}$ starting from λ_0. Keep decreasing λ_n until $Dens(\lambda_n)$ reaches the plateau and remains stable afterwards. Denote the maximum $Dens(\lambda_n)$ in its profile as $Dens_{\max}$.
- λ_{platu} =the largest λ_n s. t. $\frac{|Dens(\lambda_n) - Dens_{\max}|}{Dens_{\max}} < \varepsilon$

  : the point where $Dens(\lambda_n)$ becomes stabilized and $\boldsymbol{\Omega}$ is close to $Dens_{\max}$.
- $\lambda_p^* = \text{argmin}_{\lambda_n}\{|Dens(\lambda_n) - p \times Dens_{\max}|\}$

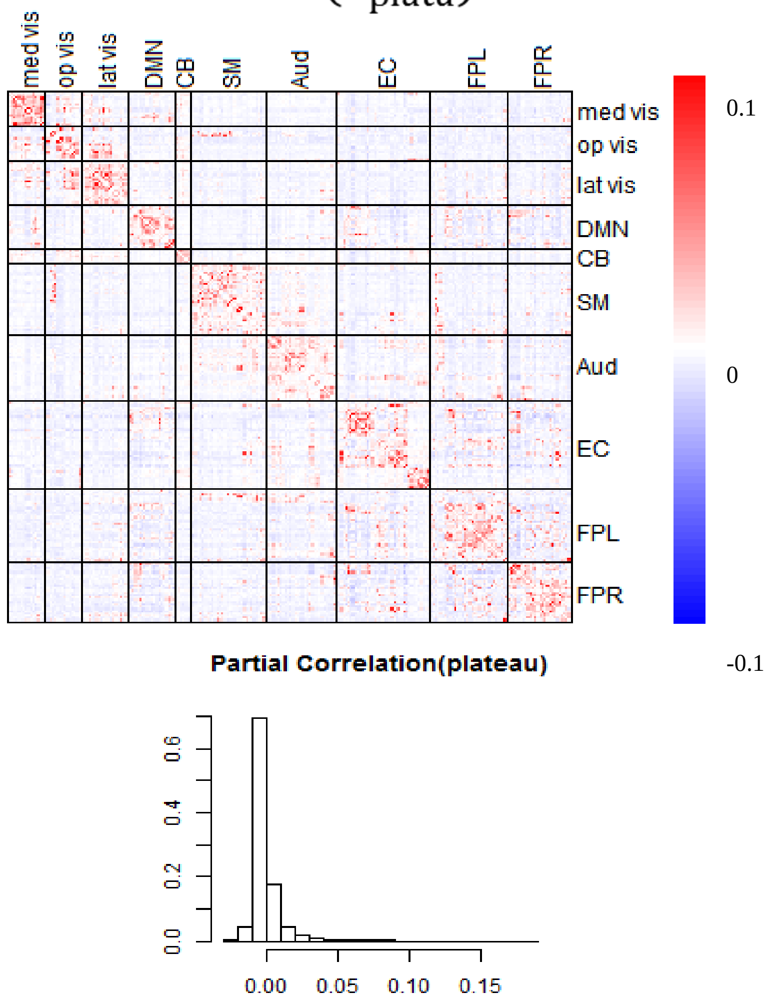
  : the point where $\boldsymbol{\Omega}$ corresponds to p percent of $Dens_{\max}$

Marginal vs. Direct connectivity

Full Correlation connectivity

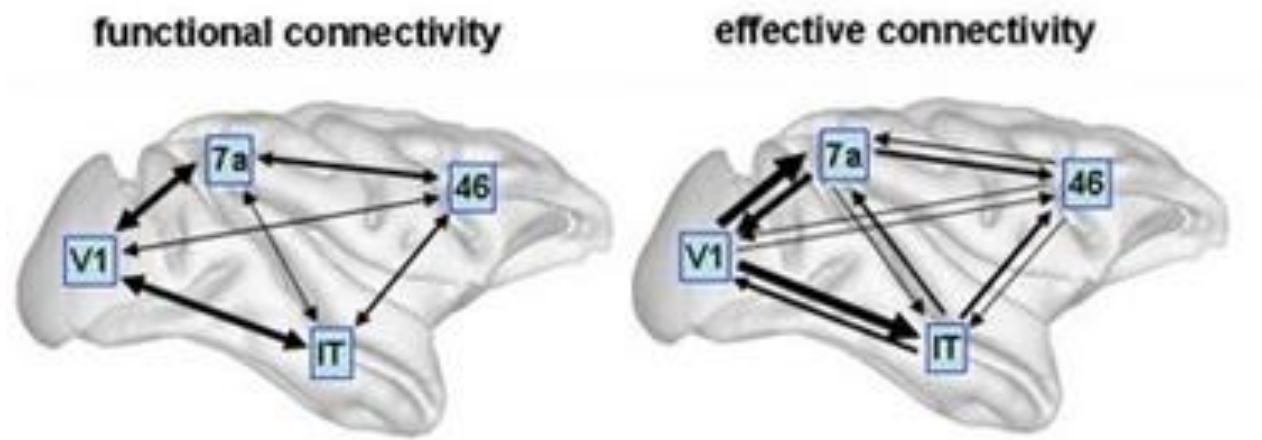


Partial Correlation connectivity (λ_{platu})



Effective Connectivity

- <u>Directed</u> influence of one brain region on the activity recorded in another brain region.



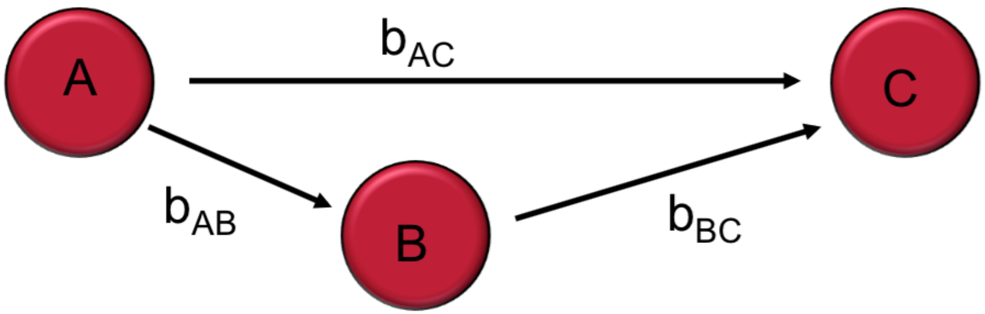
Undirected associations

Directed associations

- Methods: SEM, DCM, Granger Causality

SEM

- Structural Equation Models comprise a set of regions and a set of directed connections



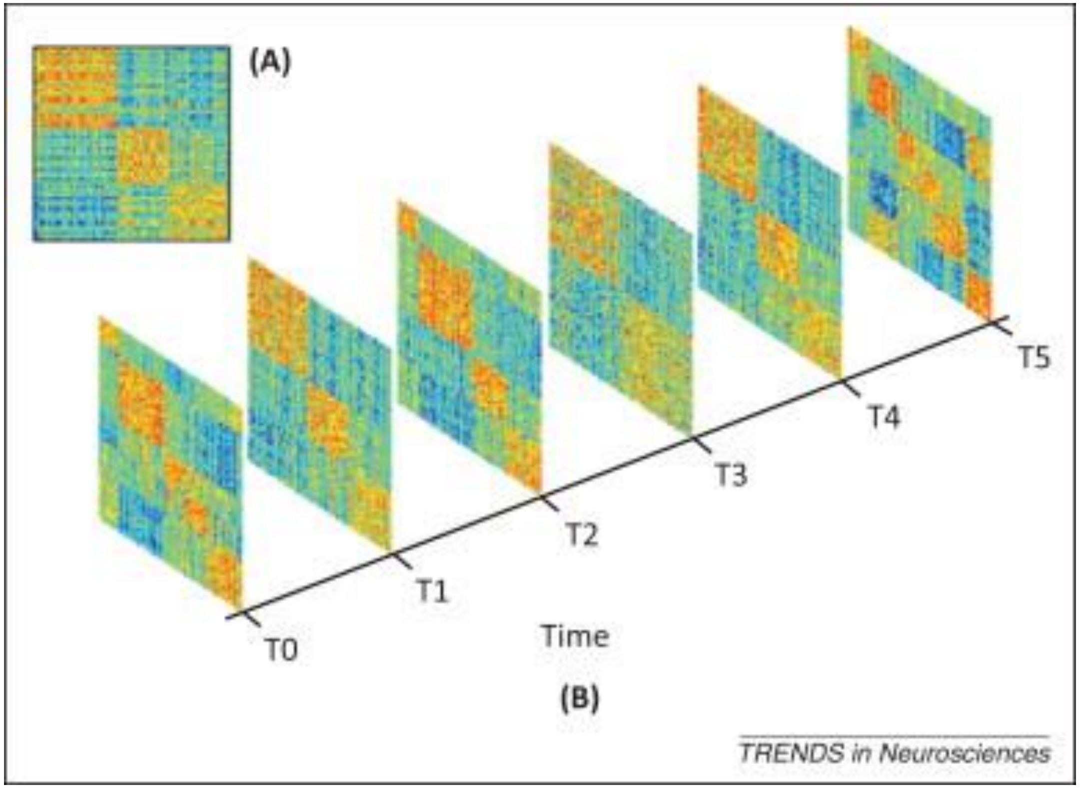
- Focuses on the covariance structure that reflects associations between variables

DCM

- Dynamic Causal Modeling estimates effective connectivity in a Bayesian framework.
- DCM regards the brain as a deterministic nonlinear dynamic system that receives inputs and that produces outputs.
- This dynamic system is modelled using neural state equation based on hemodynamic time series
- Effective Connectivity is parameterized in terms of the coupling among unobserved neuronal activity in different regions.

Dynamic FC

- Dynamic FC attempts to model changes in FC over time
- Sliding Window approach



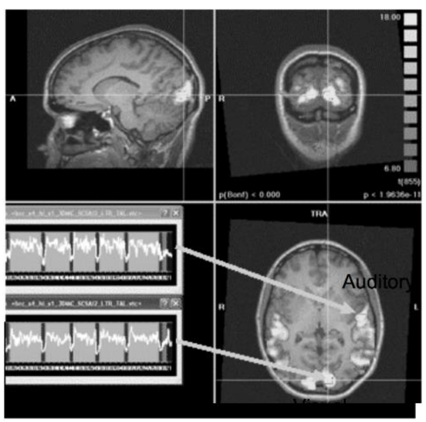
Statistical Analysis

- Goals of statistical analysis of fMRI data:
  - **ACTIVATION**: Localizing brain areas activated by the experimental task
  - **FUNCTIONAL CONNECTIVITY**: Determining networks corresponding to brain function
  - **PREDICTION**: making predictions about psychological or disease states

Prediction/Classification

- Predicting future neural activity based on baseline functional brain images.

- Predicting experimental conditions, cognitive states and group membership (psychiatric conditions, treatment response) based on functional brain images.



Experimental condition, cognitive states

Pittsburgh Brain Activity Interpretation Competition

Clinical outcomes:
•Diseased (e.g. ADHD) vs. normal
•Treatment Response vs. non-response

Prediction/Classification

- There is a growing interest in using fMRI data for classification of mental disorders and prediction of neural activity.

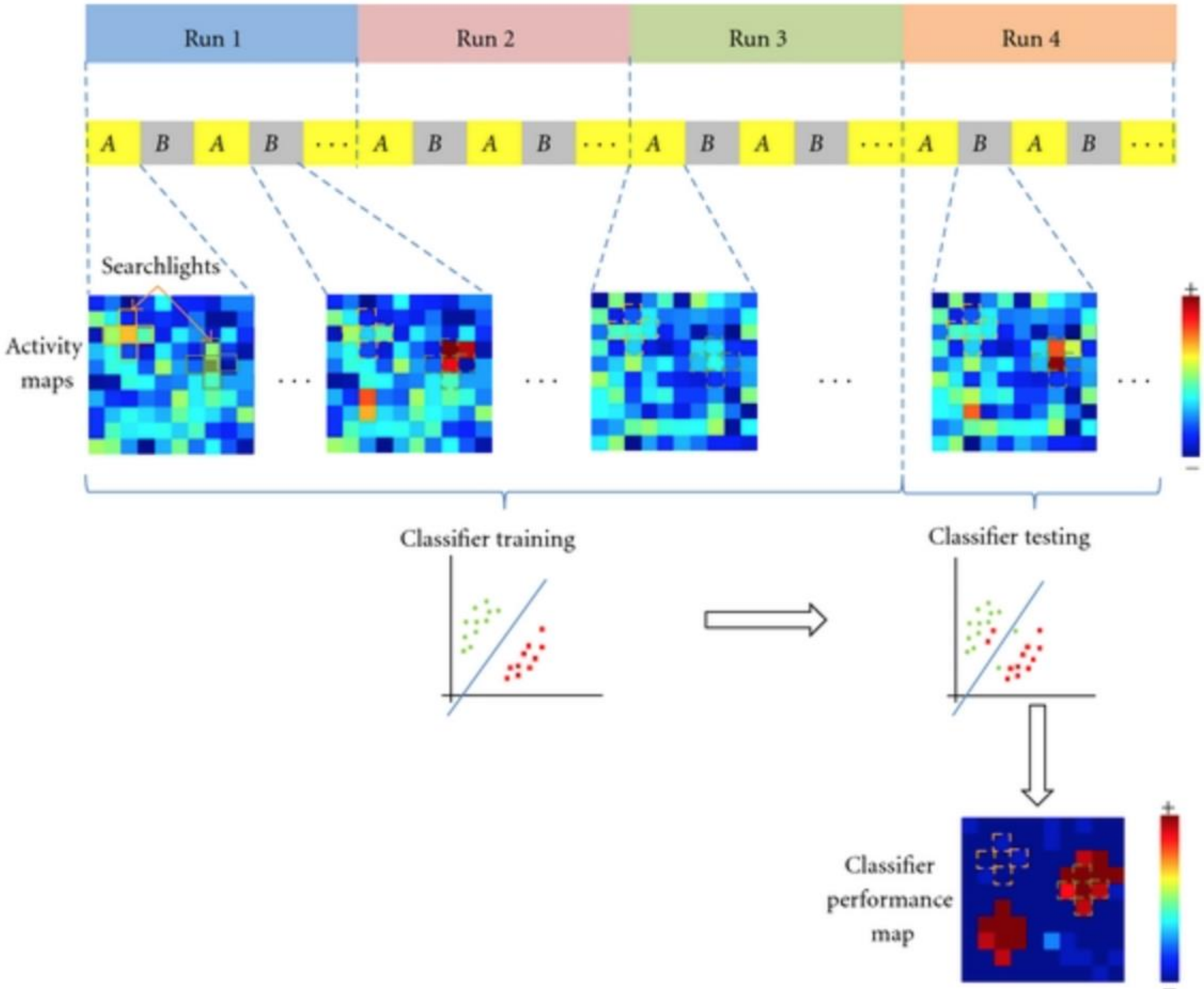
- This application of machine learning techniques is often referred to as multi-voxel pattern analysis (MVPA)
  - A classifier is trained to discriminate between different brain states and used to predict the states in a new set of data

- When applied to fMRI data, the result is often a pattern of weights across brain regions that quantify the degree to which the pattern of brain activity responds to a particular type of event. (Ex: SVM)

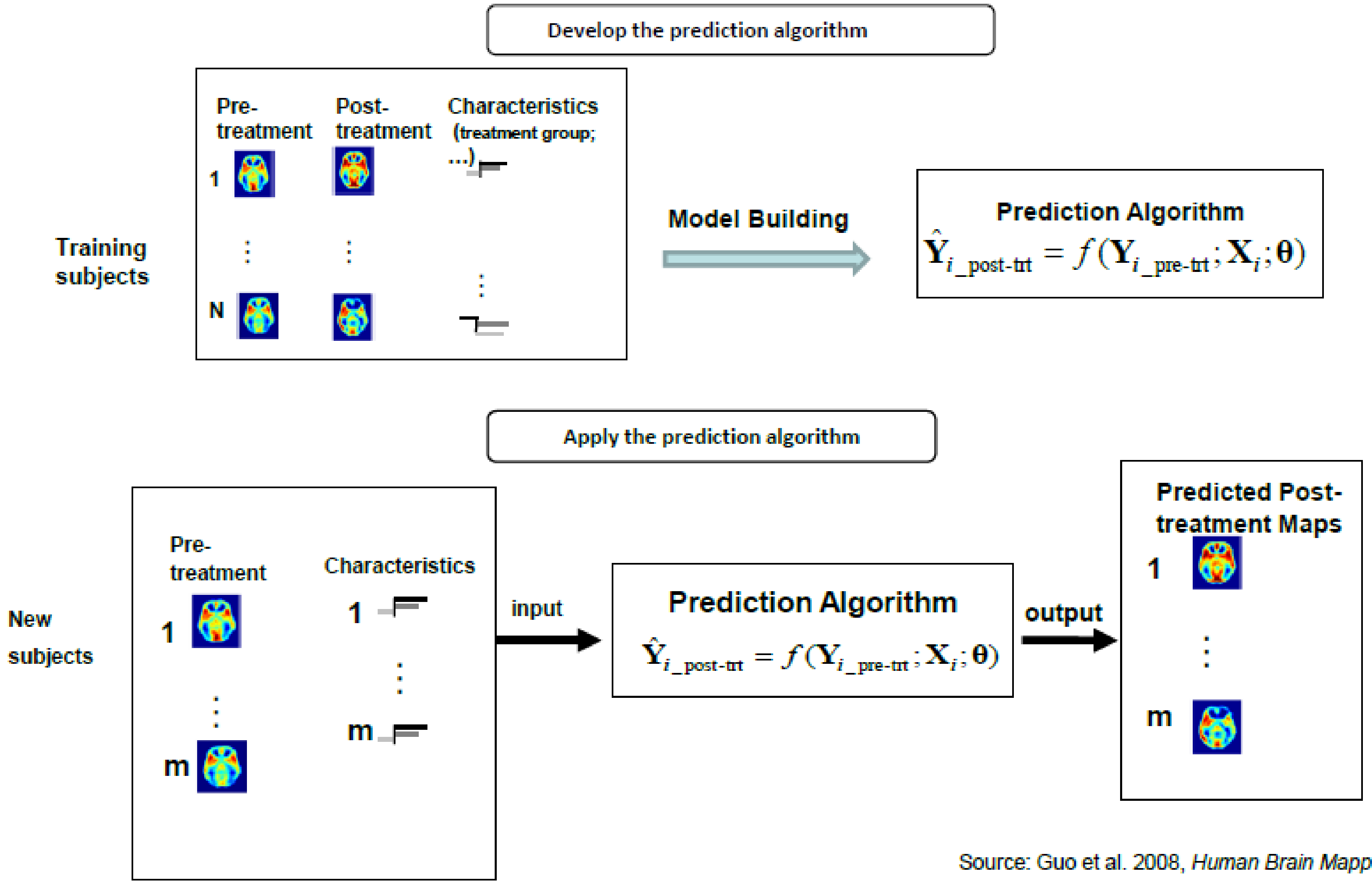
Performing MVPA

- The process of performing MVPA follows a series of steps:
  - Defining features and classes
  - Feature selection
  - Choosing a classifier
    - SVD, LDA, logistic regression
  - Training and testing the classifier
    - Cross validation
  - Examining results
    - Prediction accuracy

Performing MVPA



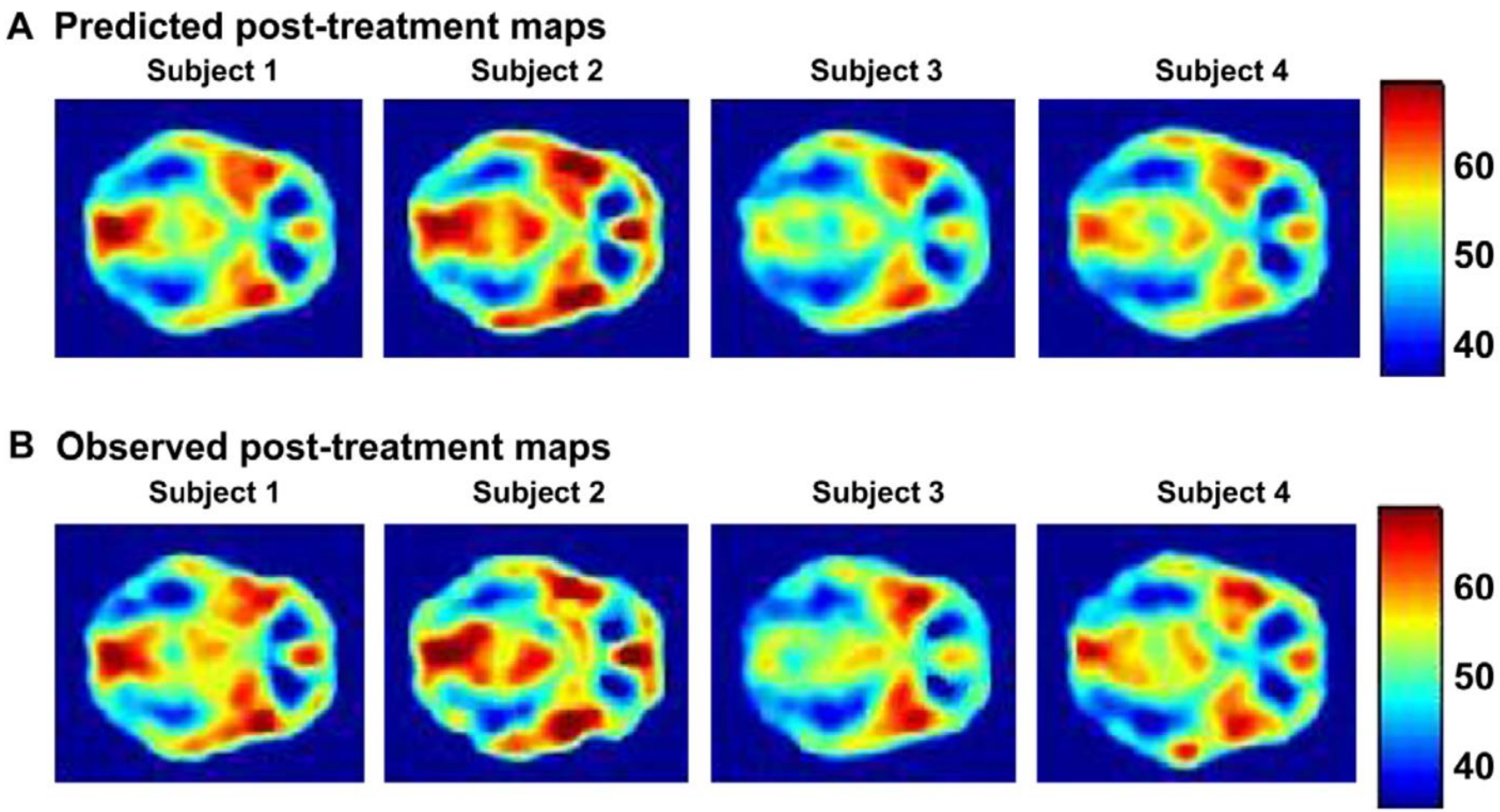
Predicting future neural activity



Source: Guo et al. 2008, *Human Brain Mapping*.

Prediction Example

Prediction of treatment response (Guo et al., 2008)



Prediction for clinical outcomes

imaging data

$\mathbf{Y}_i$

Feature selection

$\downarrow$

$\mathrm{h}(\mathbf{Y}_i)$

Building a predictive model

$\mathrm{h}(\mathbf{Y}_i) \longrightarrow f(\mathrm{h}(\mathbf{Y}_i), \mathbf{z}_i; \boldsymbol{\theta})$

non-imaging outcomes

$\mathbf{x}_i$

$\uparrow$

Prediction, training and testing

e.g.
1.Between-region functional connectivity (Shuo et al.);
2. cluster-specific principal features extracted using kernel PCA (Guo, 2010)

For categorical outcome: Support vector classifier

$$f(\mathrm{h}(\mathbf{Y}_i); \boldsymbol{\theta}) = sign[\mathrm{h}'(\mathbf{Y}_i)\boldsymbol{\beta} + \beta_0]$$

For continuous outcome: regression models such as ridge regression

$$f(\mathrm{h}(\mathbf{Y}_i); \boldsymbol{\theta}) = \beta_0 + \mathrm{h}'(\mathbf{Y}_i)\boldsymbol{\beta}$$

$$\text{with } \hat{\boldsymbol{\beta}} = \arg\min_{\boldsymbol{\beta}} \left\{ \sum_{i=1}^{N} \| \mathbf{x}_i - [\beta_0 + \mathrm{h}'(\mathbf{Y}_i)\boldsymbol{\beta}] \|^2 + \lambda \| \boldsymbol{\beta} \|^2 \right\}$$

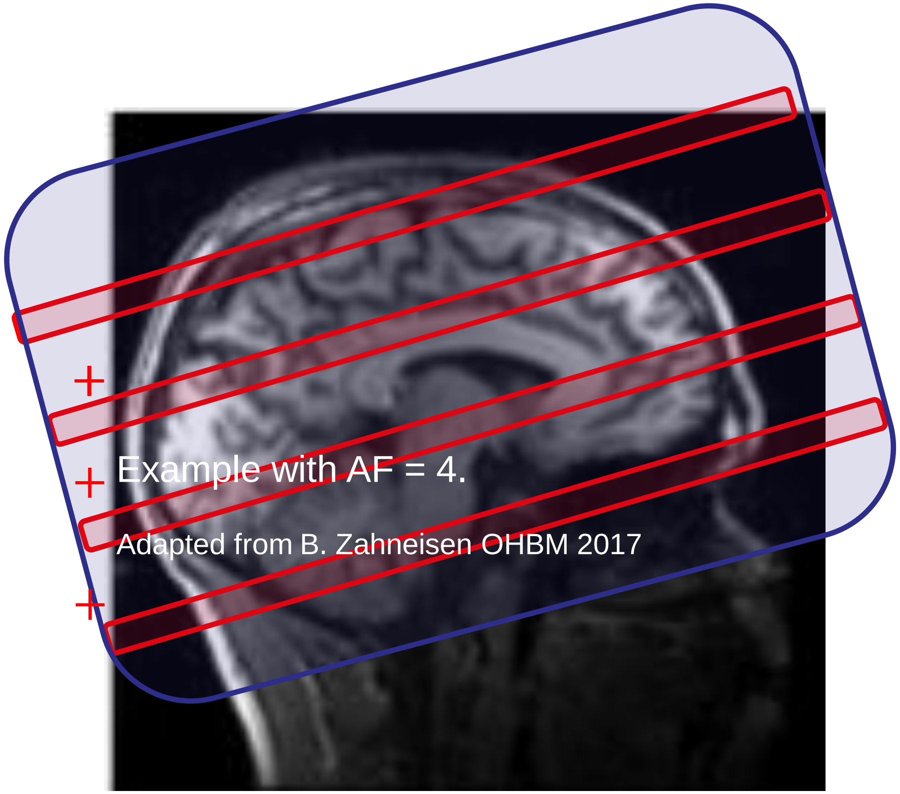
B. Risk

Collaborators: D. Rowe, M. Kociuba, J. Wu, R. Murden, D. Qiu

MULTIBAND ACQUISITION

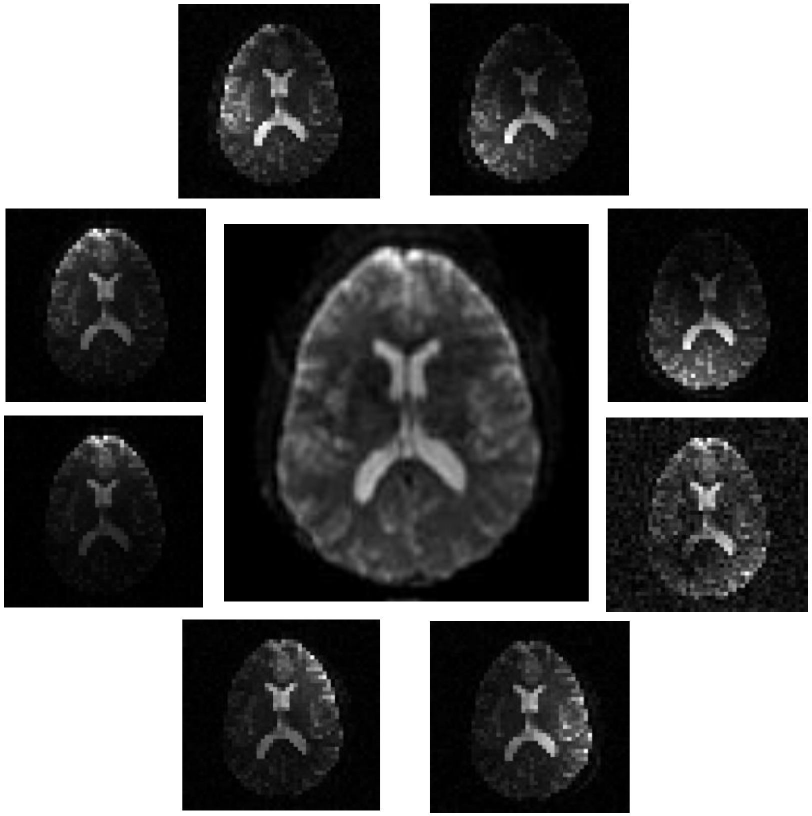
SMS Overview

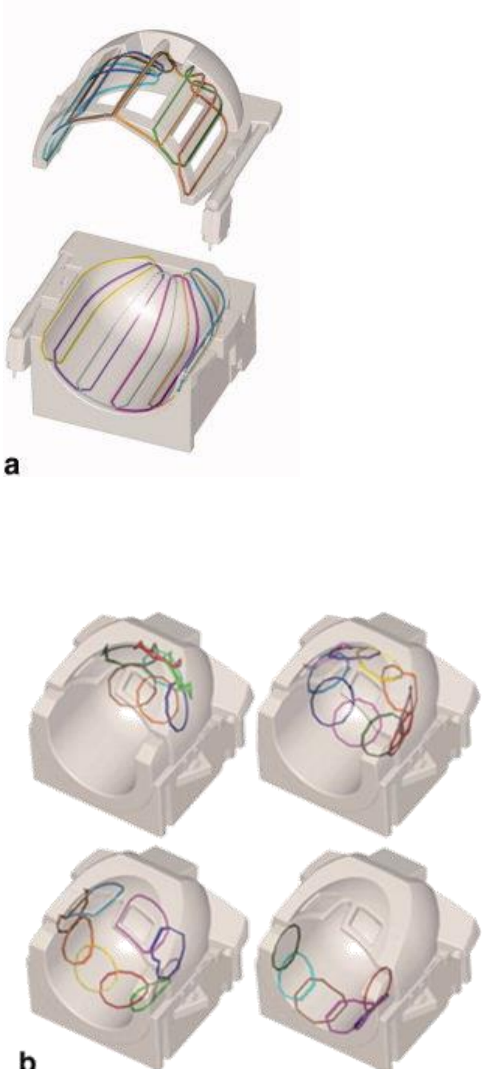
- SMS = multiband
- Multiband RF pulse with slice-selective gradient → simultaneously collect multiple slices
- Sum slices in packet
- Decrease TR
- Popular in DWI and fMRI
- Here, focus on fMRI



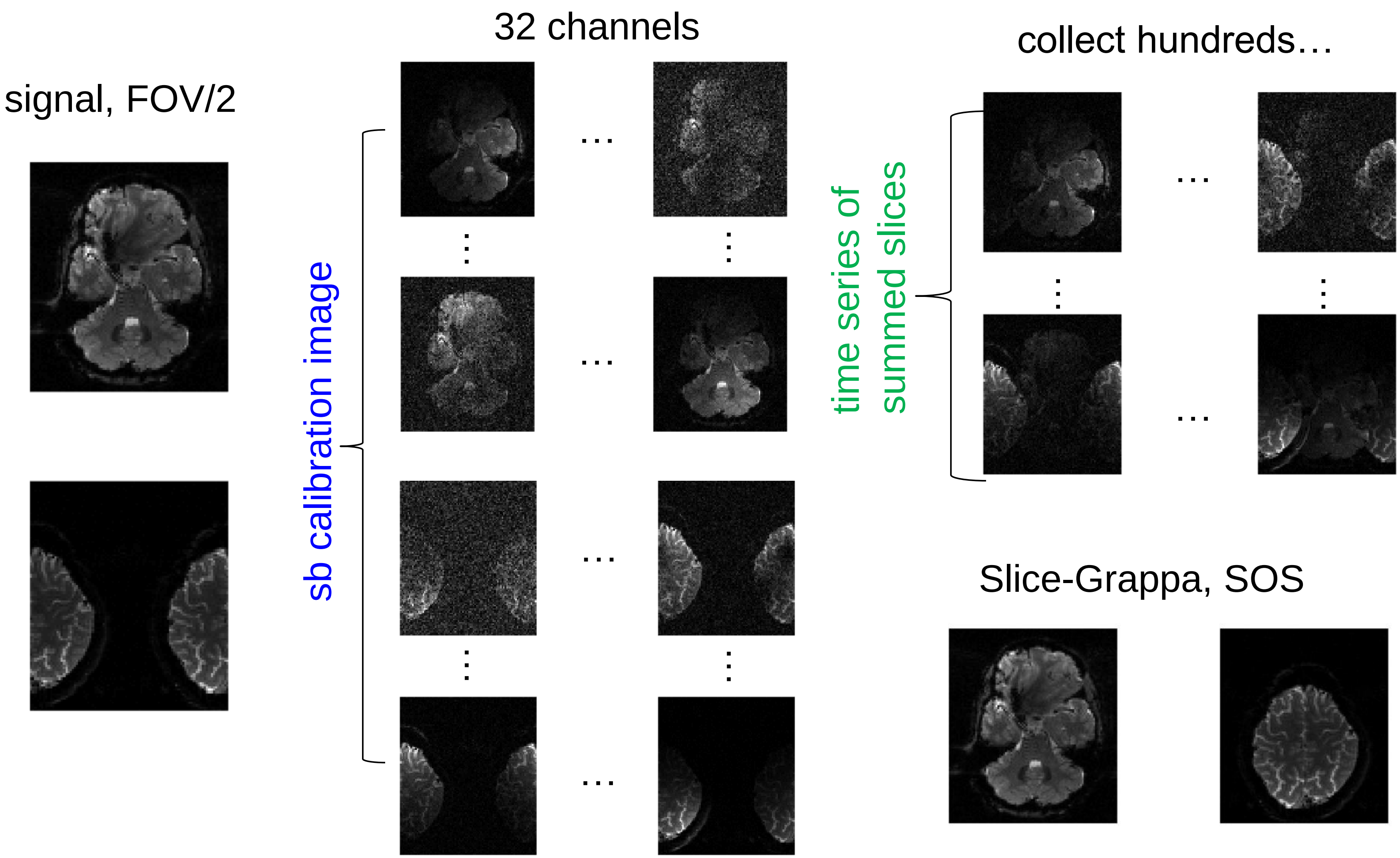
Head coil array

Coil sensitivity variation



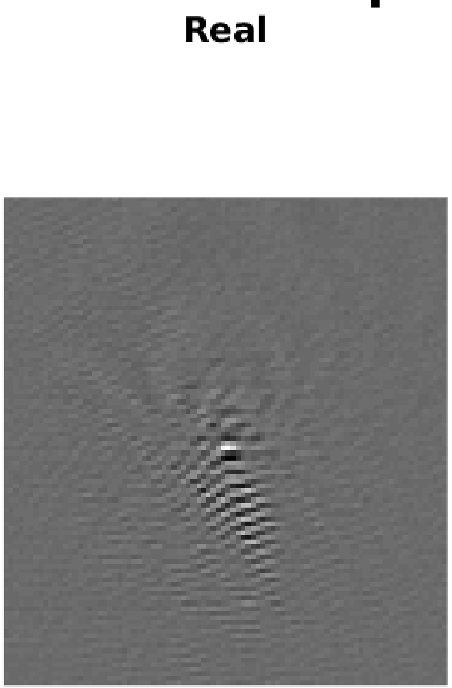


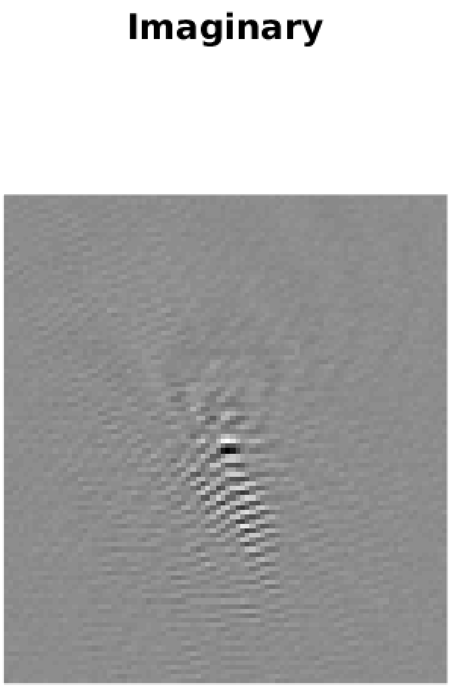
Example SMS = 2

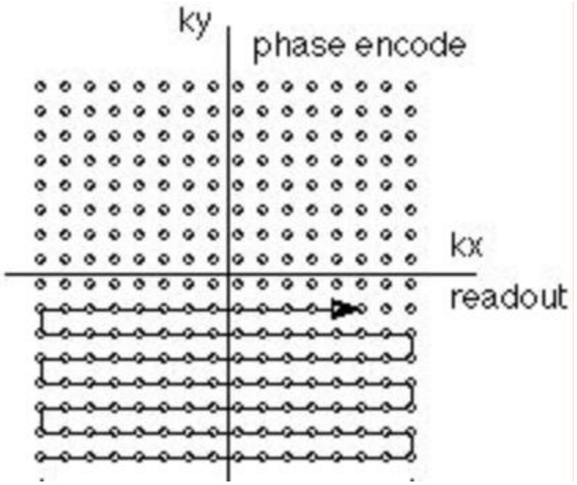


K-space

- MR data are not collected at locations
- Spatial frequencies = kspace
- 5D complex-valued data







Slice-GRAPPA

- Calibration Data (in k-space):

$$M^K_{c\ell km0} = \sum_{z\in\{m,m+M,\ldots,m+(A-1)M\}} S_{c\ell kz0}$$

$$S^K_{c\ell kz0} = \left\{\sum_{h=1}^{C}\sum_{j=-J}^{J}\sum_{i=-I}^{I} \eta_{chjiz} M^K_{h,\ell+j,k+i,m(z),0}\right\} + \epsilon_{c\ell kz0}$$

where $\text{Real}(\epsilon_{c\ell kz0}) \sim \mathcal{N}(0, \sigma^2_{cz,R})$ and $\text{Imag}(\epsilon_{c\ell kz0}) \sim \mathcal{N}(0, \sigma^2_{cz,I})$.

- Design matrix:

$$\mathbf{M}^K_{c\ell km(z)0} = [M^K_{c,\ell-J,k-I,m(z),0}, \ldots, M^K_{c,\ell+J,k+I,m(z),0}]^T \in \mathbb{C}^{(2I+1)(2J+1)}$$

$$\mathbf{M}^K_{\ell km(z)0} = [(\mathbf{M}^K_{1\ell km(z)0})^T, \ldots, (\mathbf{M}^K_{C\ell km(z)0})^T]^T \in \mathbb{C}^{C(2I+1)(2J+1)},$$

$$\mathbf{M}^K_{m(z)0} = [\mathbf{M}^K_{11m(z)0}, \ldots, \mathbf{M}^K_{YXm(z)0}]^T \in \mathbb{C}^{YX\times C(2I+1)(2J+1)}.$$

Estimate kernel

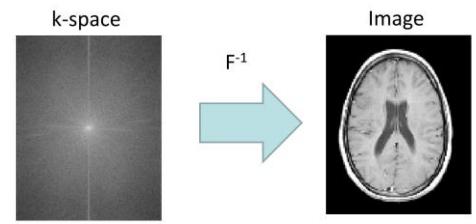
- Estimate kernel:

$$\hat{\boldsymbol{\eta}}_{cz} = ({\mathbf{M}^{K*}}_{m0}{\mathbf{M}^{K}}_{m0})^{-1}{\mathbf{M}^{K*}}_{m0}\mathbf{S}^{K}_{cz0}$$

- Apply to test data (k-space fMRI time series):

$$\hat{S}^{K}_{c\ell kzt} = \sum_{h=1}^{C}\sum_{j=-J}^{J}\sum_{i=-I}^{I}\hat{\eta}_{chjiz}M^{K}_{h,\ell+j,k+i,m(z),t}$$

- Transform to image space:



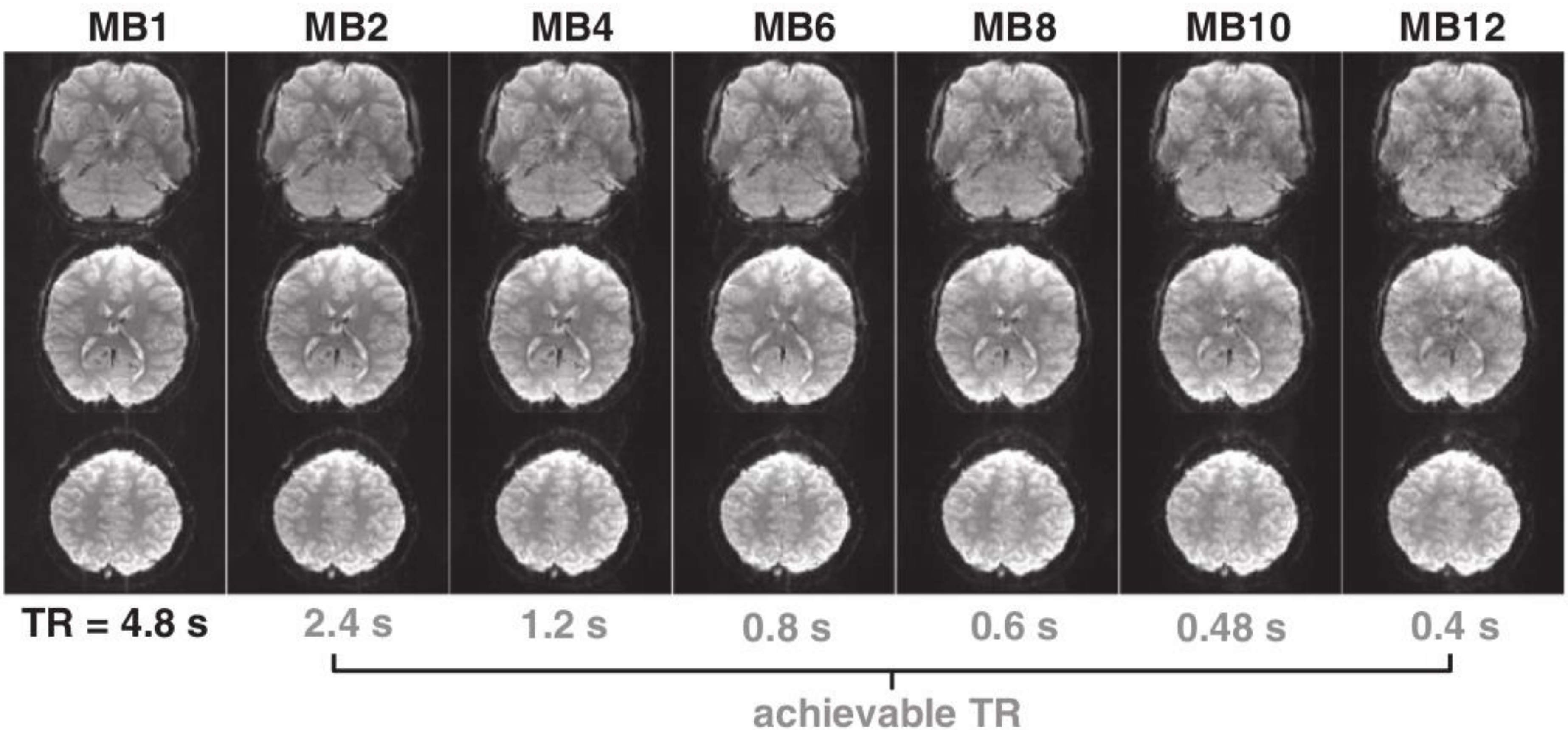
- Calculate magnitude images:

$$\hat{S}^{I}_{xyzt} = \sqrt{\sum_{c=1}^{C}\mathrm{Re}(\hat{S}^{I}_{cxyzt})^2 + \mathrm{Im}(\hat{S}^{I}_{cxyzt})^2}$$

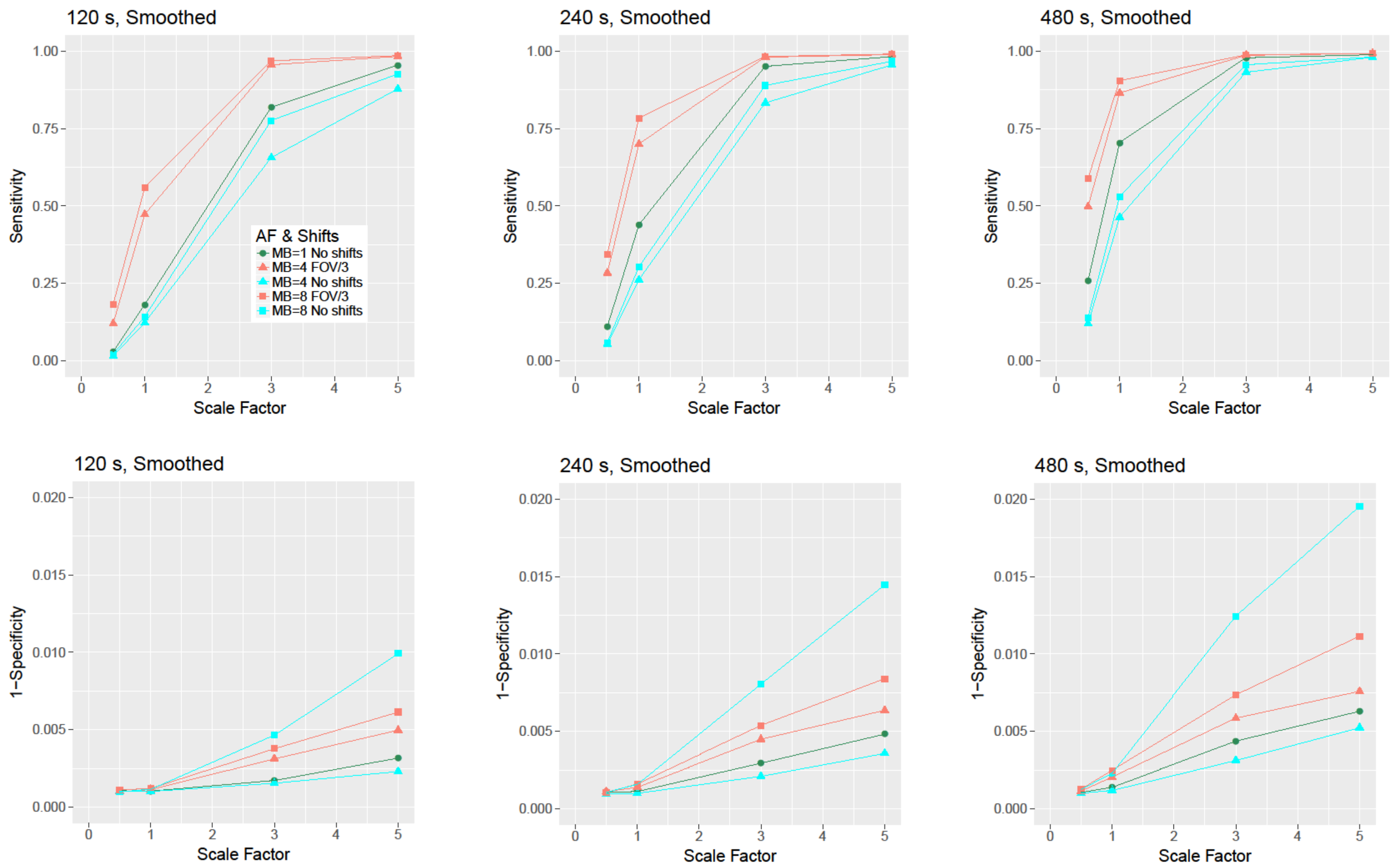
Reconstruction error

- Xu et al 2013 HCP Consortium:

Note: MB factor = SMS factor



Slice-GRAPPA



Noise amplification

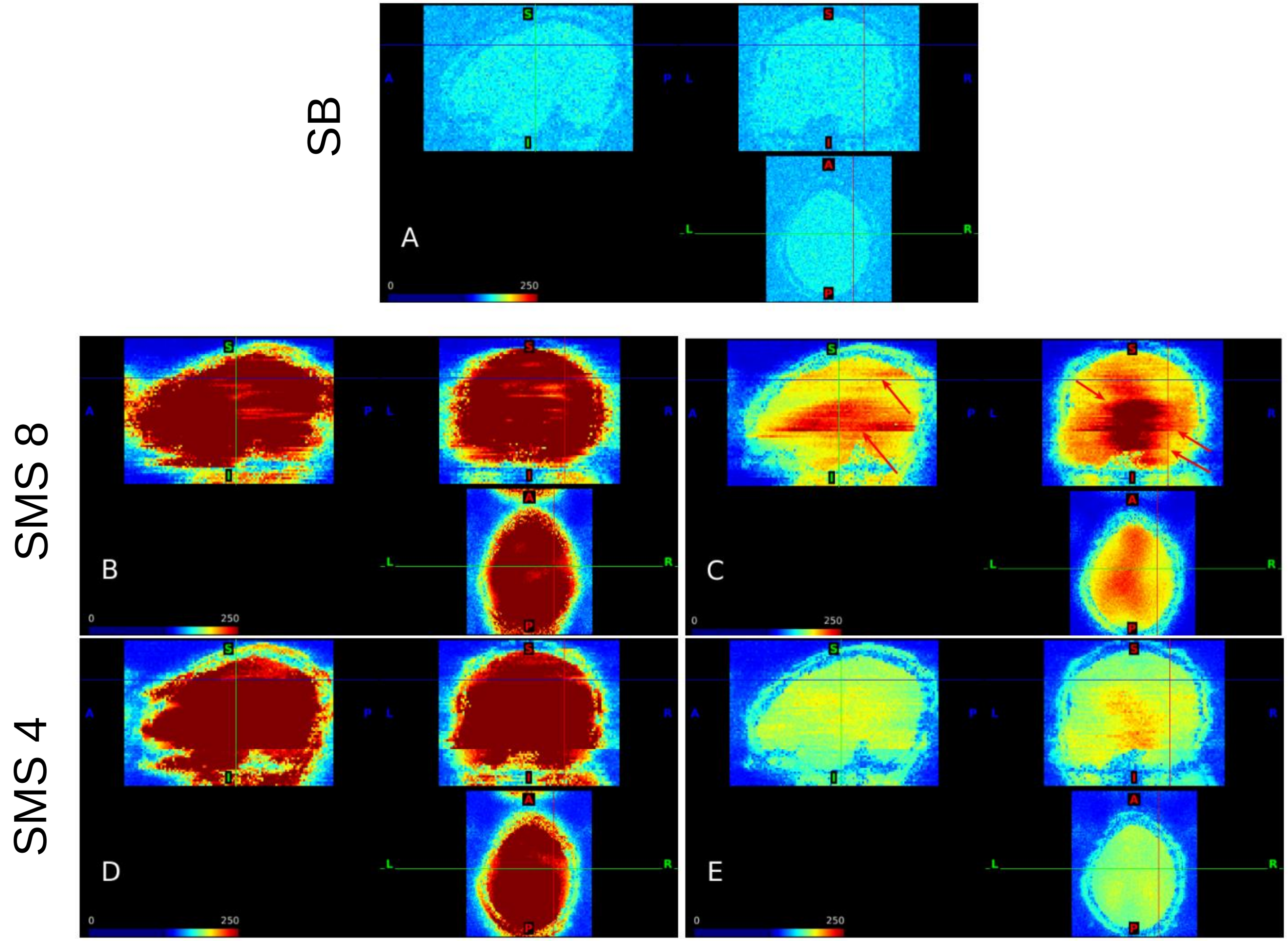
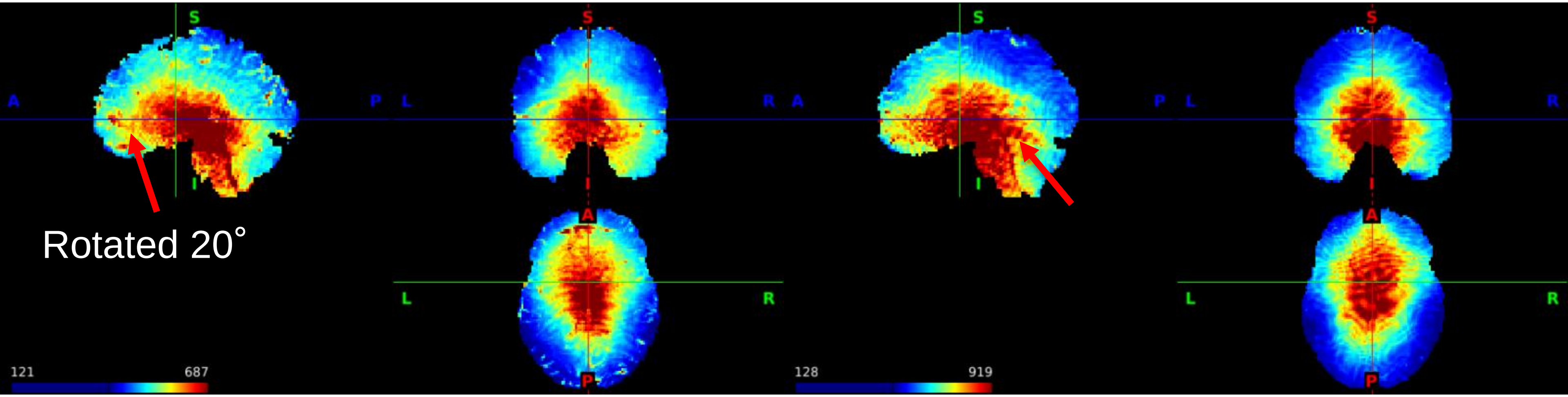
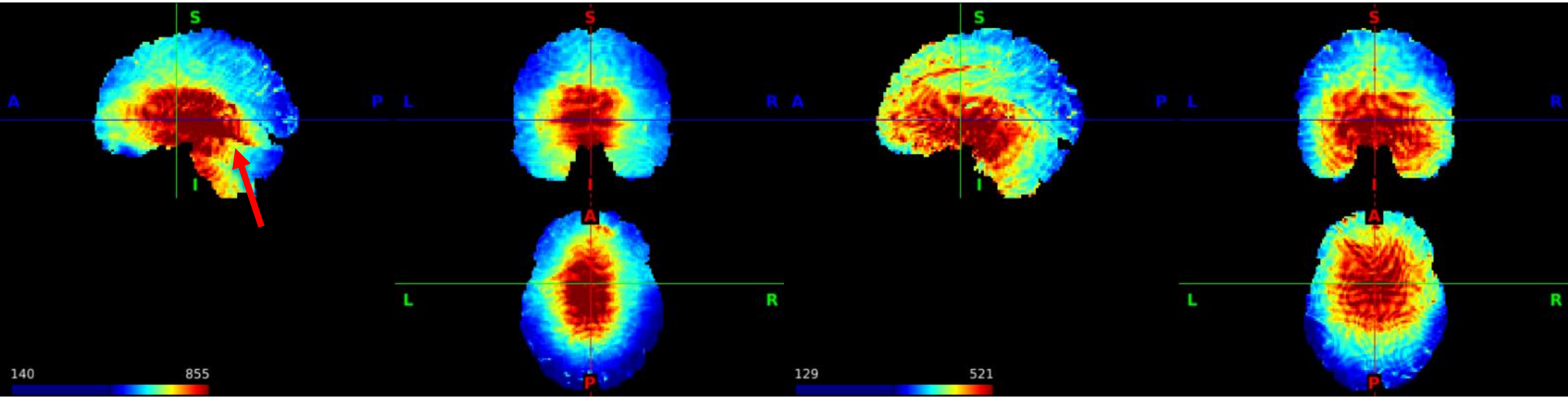


Figure 2: Noise amplification due to SMS. Standard deviation of the residuals from the GLM fit to simulations with scaling factor = 1 and scan duration = 480 s. AF = 1 (A); AF = 8 with no FOV shifts (B) and FOV/3 shifts (C); AF = 4 with no FOV shifts (D) and FOV/3 shifts (E).

SD: ICA-FIX rs-fMRI





3942. Which multiband factor should you choose for your resting-state fMRI study?

Benjamin B. Risk[a], Junjie Wu[b], and Deqiang Qiu[b,c]

[a]Department of Biostatistics and Bioinformatics, Emory University; [b]Department of Radiology and Imaging Sciences, Emory University School of Medicine

[c]Joint Department of Biomedical Engineering, Emory University and Georgia Institute of Technology

INTRODUCTION

- Thalamic and subcortical functional connectivity are useful biomarkers.[1,2]
- Multiband (MB) / Simultaneous Multislice (SMS) is used in rs-fMRI to increase temporal resolution.[3]
- Benefits of reduced TR may be decreased by noise amplification, which varies across space and is generally higher in subcortical regions.[4]

MOTIVATION

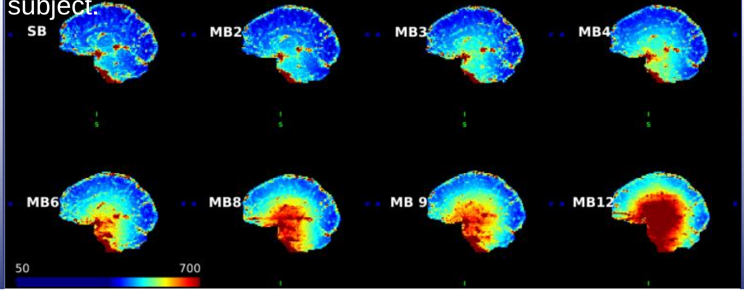
- Let $\boldsymbol{x}_{v,t,r}$ denote the BOLD signal at location v, time t, and acceleration (multiband) factor r.
- For $r' > r$ (higher AF) and v in a region of high g-factor, we hypothesize

$$Cov(\boldsymbol{x}_{v,t,r},\ \boldsymbol{x}_{v',t,r}) \approx Cov(\boldsymbol{x}_{v,t,r'},\ \boldsymbol{x}_{v',t,r'}),$$
$$sd(\boldsymbol{x}_{v,t,r}) < sd(\boldsymbol{x}_{v,t,r'}).$$

Then,

- $Corr(\boldsymbol{x}_{v,t,r},\ \boldsymbol{x}_{v',t,r}) > Corr(\boldsymbol{x}_{v,t,r'},\ \boldsymbol{x}_{v',t,r'})$.

Figure 1. Standard deviation of time series for an example subject.



Differences in functional connectivity due to noise amplification are important:

1. The magnitudes are interpreted as strength of functional connectivity, and spatially varying g-factors may mischaracterize brain activity.
2. Smaller correlations decrease statistical power unless sufficiently offset by increases in effective sample size.

Program #3942. Multiband and resting-state fMRI.

METHODS

- MR scans performed on a Siemens Prisma 3T scanner using a 32-channel head coil.
- Nine rs-fMRIs, each lasting 6 minutes, were collected on 10 subjects (5 female) ages 19-29 using either a standard or multiband EPI sequence from the Center for Magnetic Resonance Research, University of Minnesota.
- TE = 32 ms for all scans, minimal TR allowed for the respective MB factors and flip angle set to the respective Ernst angle.
- Rigid-body aligned to the single-band reference using MCFLIRT.
- Distortion correction using FSL's topup with a PA acquisition, slice time corrected, FNIRT to MNI space with the T1 image.
- 2 mm FWHM smoothing, linear+quadratic detrending, nuisance regression using global, CSF, and white matter signals and 6 parameters from motion correction.
- Notes: no temporal filtering or ICA artifact removal.
- Pearson correlation between 10-mm spheres from 264 nodes in Power atlas.[5]

MB factor	1	1	2	3	4	6	8	9	12
TR (ms)	3000	5470	2750	1850	1400	960	724	655	502
Voxel Size	3.3	2	2	2	2	2	2	2	2
Flip Angle	81	88	80	72	65	57	51	48	43

Program #3942. Multiband and resting-state fMRI.

RESULTS

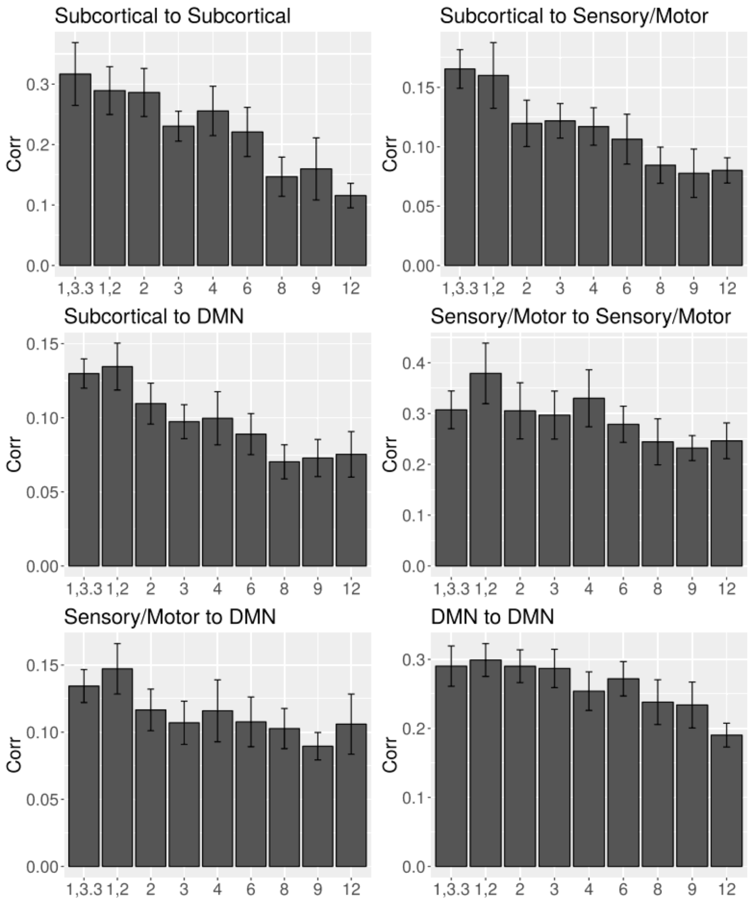


Figure 2. Correlation between edges containing nodes categorized as subcortical, sensory/motor, or default mode.

Displayed are means+/-2SE (n=10 subjects); subject correlations were calculated for each edge in the Power 264 atlas, then positive correlations from edges belonging to each pair of communities were averaged.

Using mixed models grouping by subject, MB was significant in all six models ($p<0.001$).

For subcortical-subcortical model:
Reference level: SB 2 mm ave correlation = 0.29

- SB 3.3 mm: +0.03, p=0.28
- MB 2: -0.003, p=0.90
- MB 3: -0.06, p=0.02
- MB 4: -0.03, p=0.19
- MB 6: -0.07, p<0.01
- MB 8: -0.14, p<0.0001
- MB 9: -0.12, p<0.0001
- MB 12: -0.17, p<0.0001

Program #3942. Multiband and resting-state fMRI.

RESULTS

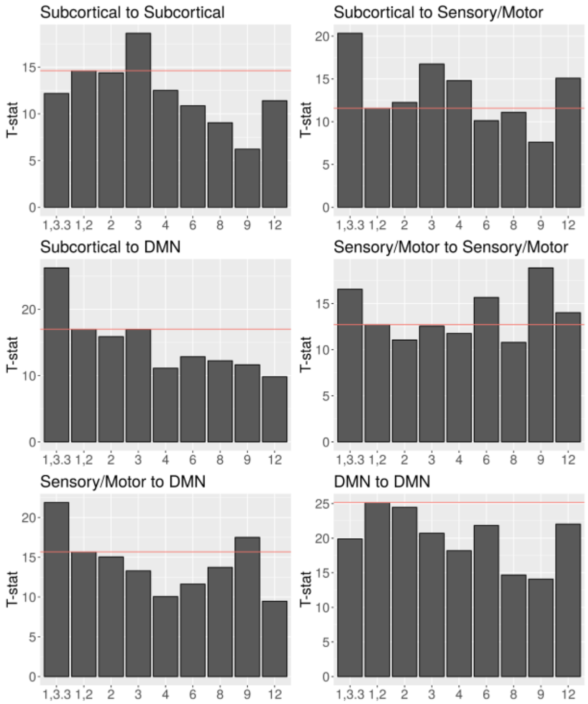


Figure 3. Preliminary optimality results. T-statistics for H_0: $Corr\left(\boldsymbol{x}_{v,t,r}, \boldsymbol{x}_{v',t,r}\right) = 0$, based on Figure 2 (n=10 subjects). T-stat larger than T-stat SB 2 mm (red line) is evidence of net benefits.

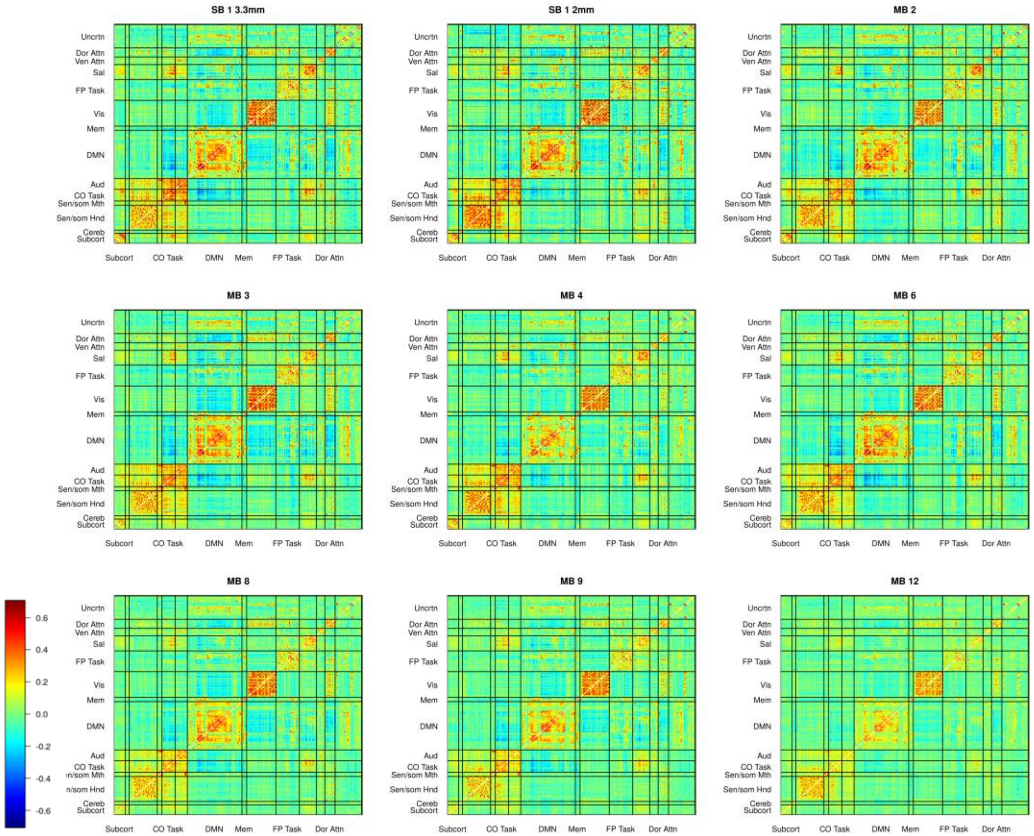
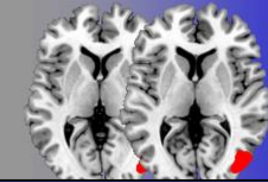


Figure 4. Correlations averaged across subjects. Note correlations with subcortical regions decrease with multiband factor (lower left corner), while visual regions tend to be less sensitive.



Program #3942. Multiband and resting-state fMRI.

DISCUSSION

- The optimal MB factor in a resting-state fMRI study depends on the regions of interest and should be considered during the design of the study.
- Results from 10 subjects clearly indicate correlations decrease with higher MB, particularly in subcortical regions.
- It would be inaccurate to make functional connectivity comparisons between regions. Example: functional connectivity in DMN-DMN versus subcortical-subcortical: SB 2mm: 0.30 vs 0.29; MB 8 2mm: 0.24 vs 0.15.
- Preliminary results suggest MB=3 in subcortical, which balances larger number of timepoints with smaller correlations.
- However, with n=10 subjects, assessment of optimality is noisy.
- Large subject heterogeneity: accuracy of distortion correction and registration, individual differences in g-factors, motion.
- Did not use ICA artifact removal or temporal filtering, which could increase MB benefits.[3,4,6]

ACKNOWLEDGMENTS

This study was funded by the Center for Systems Imaging at the Emory University School of Medicine.

FINANCIAL DISCLOSURES

None.

REFERENCES

1. Barron DS, Fox PT, Pardoe H, et al. Thalamic functional connectivity predicts seizure laterality in individual TLE patients: Application of a biomarker development strategy. *NeuroImage Clin*. 2015.
2. Kim K, Kim S-W, Myung W, et al. Reduced orbitofrontal-thalamic functional connectivity related to suicidal ideation in patients with major depressive disorder. *Sci Rep*. 2017;7(1):15772.
3. Smith SM, Beckmann CF, Andersson J, et al. Resting-state {fMRI} in the human connectome project. *Neuroimage*. 2013;80:144-168.
4. Todd N, Josephs O, Zeidman P, Flandin G, Moeller S, Weiskopf N. Functional sensitivity of 2D simultaneous multi-slice echo-planar imaging: effects of acceleration on g-factor and physiological noise. *Front Neurosci*. 2017;11.
5. Power JD, Cohen AL, Nelson SM, et al. Functional network organization of the human brain. *Neuron*. 2011;72(4):665-678.
6. Jahanian H, Holdsworth S, Christen T, et al. Advantages of Short Repetition Time Resting-State Functional MRI Enabled by Simultaneous Multi-slice Imaging. *J Neurosci Methods*. 2018.