Statistical Parametric Mapping, Multivariate Analysis and Brain Connectivity

Will Penny

Wellcome Trust Centre for Neuroimaging.
University College London.

Workshop on Statistics, Machine Learning and Neuroscience, International Conference on Machine Learning (ICML).
29th June 2012, Edinburgh.

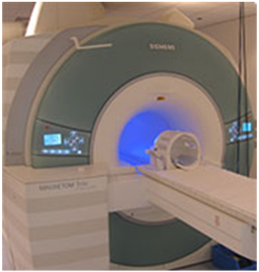
Statistical Parametric Mapping

Given data $Y = \{y_1, y_2, ...y_S\}$ comprising S time series recorded over space S, fit a linear model to each voxel i

$$
\begin{aligned}
y_i &= Xw_i + e_i \\
p(y_i|w_i) &= \mathrm{N}(y_i; Xw_i, C_i)
\end{aligned}
$$

using Maximum Likelihood.

MRI



MEG

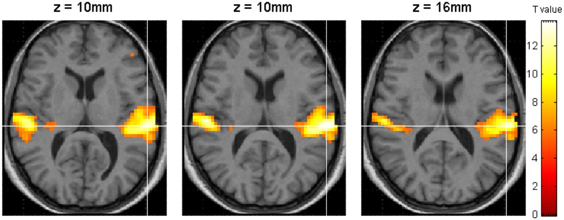
This produces images of regression coefficients $W = [w_1, w_2, ..., w_S]$ and error fields $E = [e_1, e_2, ..., e_S]$.

Statistical Parametric Mapping

Compute a ‘contrast’ of regression coefficients

$$c_i = C_i^T w_i$$

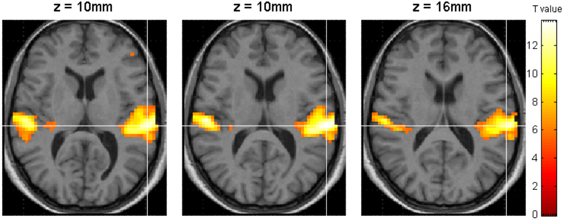
and associated images t_i or F_i statistics.



The resulting images are Statistical Parametric Maps (*Friston et al., Human Br. Mapp. 1995*).

Statistical Parametric Mapping

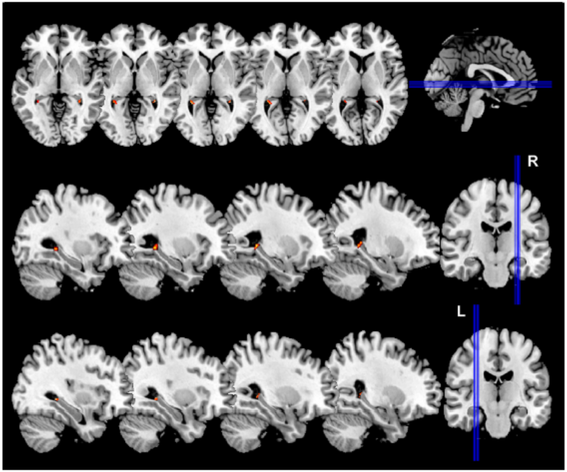
The maps are thresholded so that the probability of one of more false positives anywhere in the search space (a 'familiy wise error') is less than 5 per cent.



The threshold is usually computed using Random Field Theory based on estimates of how smooth the error fields are (*Worsley et al. Chance, 2005*).

Statistical Parametric Mapping

SPM showing regions of significant increase in Grey Matter Volume in taxi drivers after acquiring 'the knowledge' of London (versus before)



The increases are in posterior hippocampus (*Woollet and Maguire, Current Biology, 2005*).

Nested Oscillations

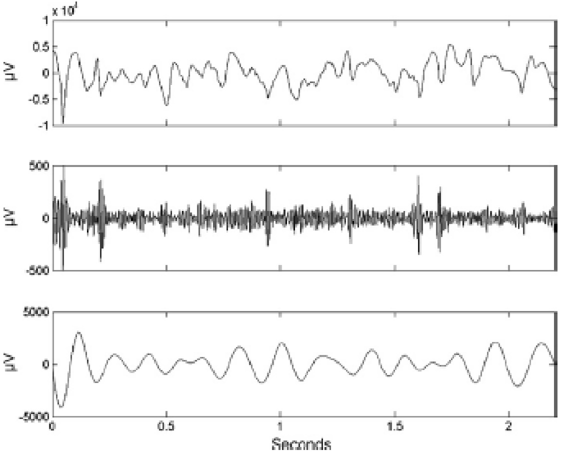


Fig. 3. ECOG time series for non-target trial. Example non-target trial at electrode 63 showing the original time series (top), activity in the χ-band (middle) and activity in the theta band (bottom). The PAC measures are $r_{ESC} = -0.42$, $r_{GLM} = 0.32$, PLV = 0.57 and $M = 12.7$.

Current item needs to be remembered.

ECoG Data

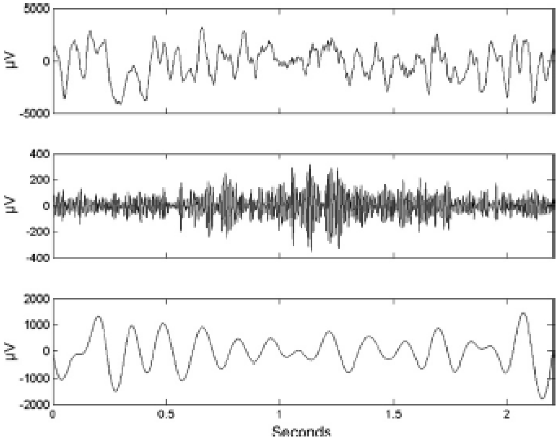


Fig. 4. ECOG time series for target trial. Example target trial at electrode 63 showing the original time series (top), activity in the χ-band (middle) and activity in the theta band (bottom). The PAC measures are $r_{ESC} = 0.02$, $r_{GLM} = 0.06$, $PLV = 0.07$ and $M = 6.8$.

Current item does not need to be remembered.

General Linear Model

Penny, J. Neurosci Methods, 2008 use a General Linear Model (GLM) approach based on the multiple regression model

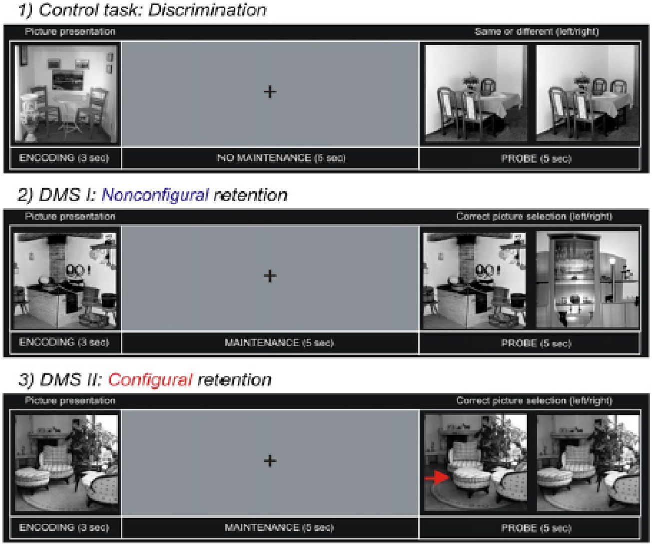
$$a_\gamma = X\beta + e$$

where β are regression coefficients, e is additive Gaussian noise and the design matrix X contains three columns:

- $\cos(\phi_\theta[n])$
- $\sin(\phi_\theta[n])$
- A column of 1's

Significance is assessed using F-tests over the first two regression coefficients. More generally, X could be a Fourier series.

MEG Experiment



MEG Study of Visual Working Memory (*Fuentemilla et al. Current Biology, 2010*).

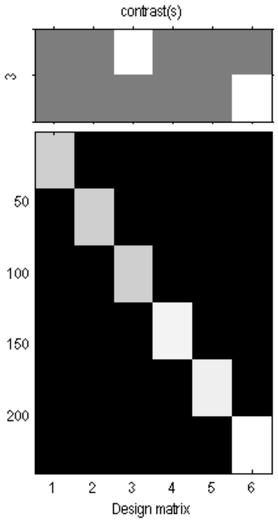
Processing Stream

- Extract phase of theta activity in source region.
- Extract time-frequency maps at each sensor, v, from frequencies $f = 16 : 4 : 128$ Hz during delay period.
- For each trial computevregression coefficients s_{fv} and c_{fv}. The sine and cosine terms for each frequency and sensor
- Create an image of regression coefficients for each trial. This then becomes data for a between-trial analysis.
- There are 3 conditions (Control, Non-configural, Configural) and 40 trials per condition, with 2 measures per trial. This gives 240 data points per subject
- Set up design matrix in SPM and implement a 'space-frequency' analysis *Litvak et al, Comput. Intell., 2010*

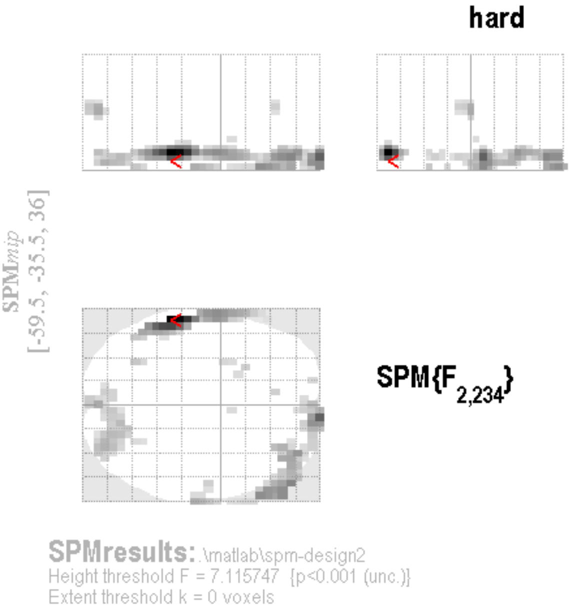
Configural

Images are entered in the following order

- Sine coefficients for Control
- Sine coefficients for Non-Config
- Sine coefficients for Config
- Cos coefficients for Control
- Cos coefficients for Non-Config
- Cos coefficients for Config



SPM


hard
SPMmip [-59.5, -35.5, 36]
SPM{F2,234}
SPMresults: .\matlab\spm-design2
Height threshold F = 7.115747 {p<0.001 (unc.)}
Extent threshold k = 0 voxels

Configural

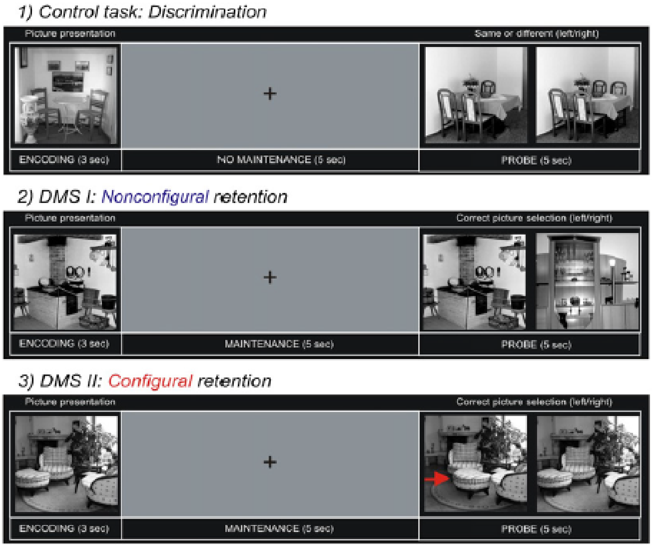
The statistical significance of phase amplitude coupling is corrected for the multiple comparisons over space and frequency using Random Field Theory.

Statistics: ***p-values adjusted for search volume***

set-level		cluster-level				peak-level					mm mm Hz		
p	c	$p_{FWE-corr}$	$q_{FDR-corr}$	k_E	p_{uncorr}	$p_{FWE-corr}$	$q_{FDR-corr}$	F	$(Z_\equiv)$	p_{uncorr}			
0.009	**12**			**77**		**0.000**	**0.003**	**19.04**	**5.48**	**0.000**	**-60**	**-36**	**44**
						0.201	0.177	11.33	4.11	0.000	-64	-3	44
				115		**0.033**	**0.079**	**13.77**	**4.59**	**0.000**	**9**	**72**	**40**
						0.134	0.177	11.90	4.22	0.000	55	51	40
						0.195	0.177	11.38	4.12	0.000	64	45	40
				12		**0.460**	**0.321**	**10.02**	**3.82**	**0.000**	**0**	**-89**	**80**
				8		**0.541**	**0.365**	**9.71**	**3.75**	**0.000**	**68**	**-19**	**32**
						0.944	0.646	7.96	3.32	0.000	55	-9	32
				52		**0.577**	**0.367**	**9.58**	**3.72**	**0.000**	**30**	**-84**	**36**
						0.706	0.479	9.11	3.61	0.000	13	-73	40
						0.760	0.515	8.91	3.56	0.000	30	-89	44
				2		**0.878**	**0.646**	**8.39**	**3.43**	**0.000**	**-30**	**-62**	**32**
				2		**0.923**	**0.646**	**8.12**	**3.36**	**0.000**	**4**	**61**	**44**
				3		**0.929**	**0.646**	**8.08**	**3.35**	**0.000**	**-26**	**18**	**40**
				3		**0.939**	**0.646**	**8.00**	**3.33**	**0.000**	**-9**	**40**	**84**
				1		**0.976**	**0.775**	**7.61**	**3.23**	**0.001**	**-17**	**67**	**32**
				3		**0.978**	**0.775**	**7.59**	**3.22**	**0.001**	**-30**	**2**	**44**
				1		**0.989**	**0.841**	**7.36**	**3.16**	**0.001**	**4**	**8**	**56**

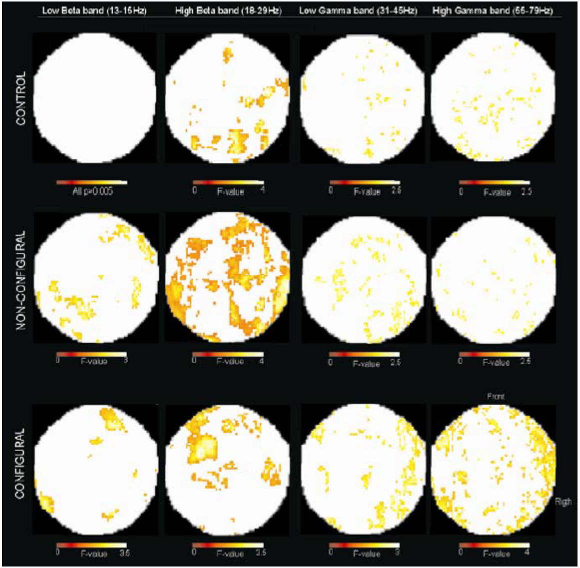
We can use the standard threshold eg FWE=0.05.

MEG Experiment



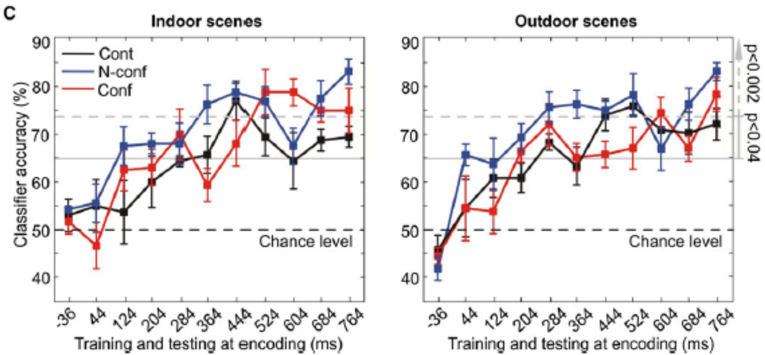
Visual scenes were either indoor or outdoor.

Multivariate Analysis at Encoding



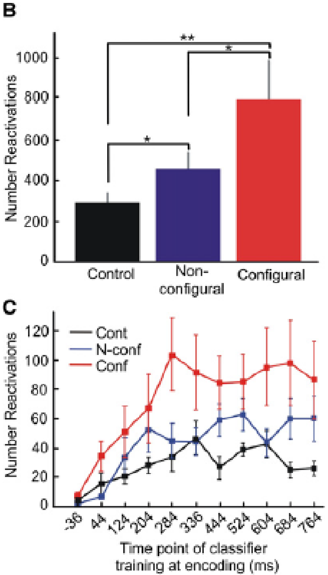
Features that discriminate between indoor and outdoor scenes.

Multivariate Analysis at Encoding



Multivariate classification (indoor versus outdoor) based on sensor space spectra using features from 13 to 80 Hz and a Multilayer Percepton with four hidden units.

Multivariate Classification of Maintenance

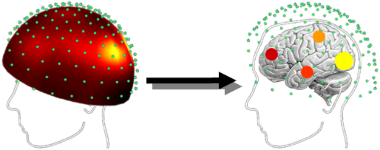


Greater replay during memory conditions.

Source Reconstruction

How do these replays relate to theta ?

Theta activity was then projected to source space *Poch et al. J. Neurosci, 2011*.



Source reconstruction is often implemented using Bayesian inference.

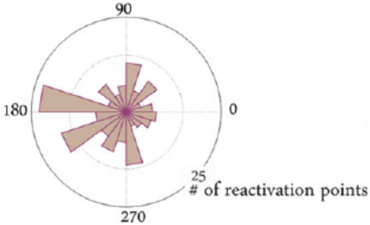
$$
\begin{aligned}
p(y|w) &= \mathrm{N}(y; Xw, C_y) \\
p(w) &= \mathrm{N}(w; m_w, C_w)
\end{aligned}
$$

Priors are needed as we may have more sources than sensors.

Replay is Phase-Locked to Theta

Further analysis by *Poch et al. J. Neurosci, 2011*

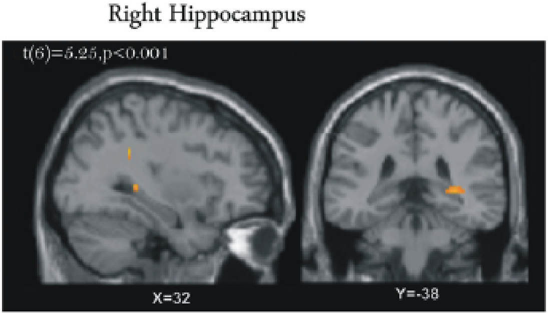
- Compute phases at which patterns were replayed.
- Assess degree of phase uniformity using PLV measure.



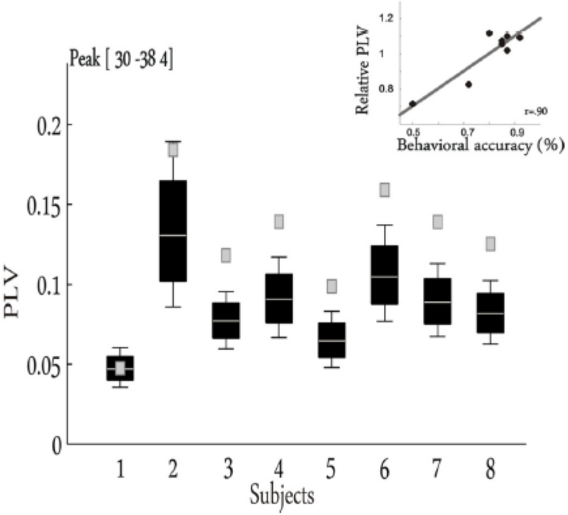
- Test for correlation between memory performance and PLV.
- SPM analysis in source space.

Theta Sources

This identified a right hippocampal and a right inferior frontal region.

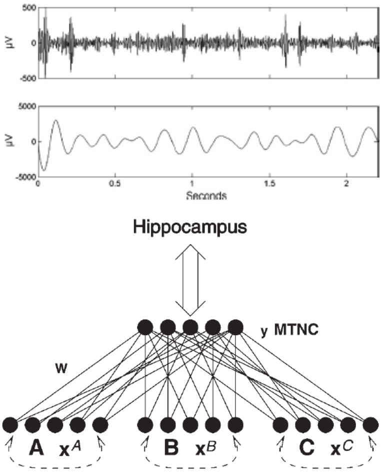


Theta Sources



Subjects with pattern reactivation times more closely coupled with hippocampal theta had better memories.

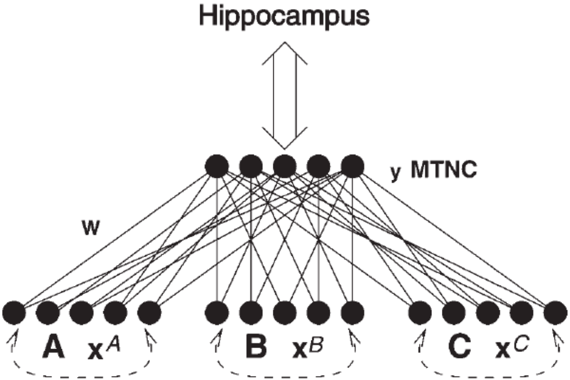
Replay in Maintenance Period



For short term memory. This replay at theta (4-8Hz).

Consolidation through replay

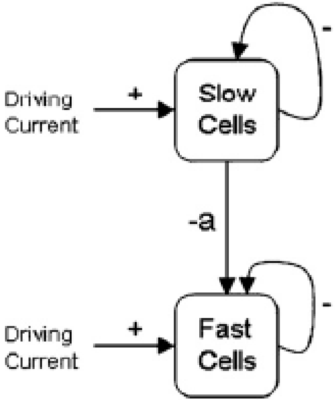
Replay during sleep allows for strengthening of new configurations of ensembles, and integration of new patterns with old (*Kali and Dayan, Nat. Neuro, 2004*)



Sleep replay is ten times faster (*McNaughton, AI, 2009*)

Hippocampal Interneuron Network

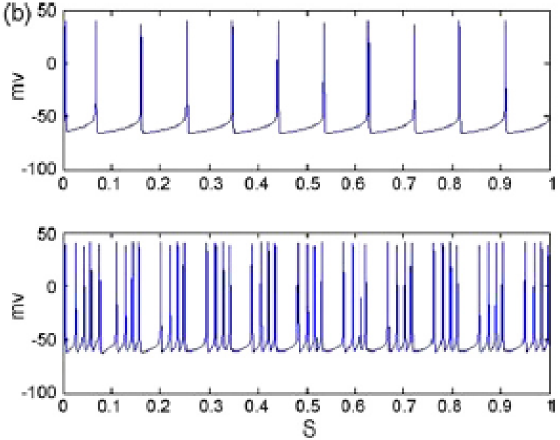
A population of Slow GABA-A cells inhibits a population of Fast GABA-A cells.



Each cell is a single compartment Hodgkin-Huxley model (*White et al, PNAS, 2000*).

Hippocampal Interneuron Network

Populations of GABA-B (top,slow) and GABA-A (bottom,fast) cells.



Fast cells had synaptic rise times of 1ms and fall times of 9ms. For the slow cells they are 5ms and 150ms.