# Advanced Signal Processing II (aka Acronym Day)

Latency Jitter and Woody Filters (acronym free)
PCA
ICA

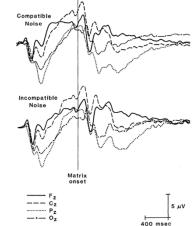
Removal of OCULAR artifacts with ICA (and lots of acronyms) BESA

Simultaneous EEG with ICA and fMRI!

# The Problem of Latency Jitter

- > The averaging assumption of invariance in signal is not always warranted
  - > Especially for the later endogenous components
  - > To the extent that the signal varies from trial to trial, the average will produce potentially misleading results
- > Two common possibilities:
  - Smearing of components;
    - will underestimate amplitude of component (especially a problem if comparing groups, one group with more latency jitter)
  - > Bimodal or multi-bumped components

# A No noise ###### ##### #RIGHT ##### ##### ##LEFT ##### #### a Noise NRIGHT KWSMNT BMJUKM UYRMUD EQEIKM VTFMZS KEHEHG ILEFTA



# The Solution

- ➤ The Woody Adaptive Filter (Woody, 1967)
- ➤ Based on Cross-correlation
  - Assumptions less restrictive than averaging methods
    - ➤ Waveform (morphology) must be constant across trials
    - ➤ Latency need not be constant

# Details

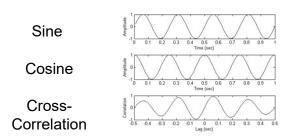
- Cross-correlational series
  - > For two waveforms the correlation between each of them is computed
    - ➤ first with no lag in time a1, a2, ..., an b1, b2, ... bn
    - then with one lagged with respect to the other a1, a2, ..., an-1 b2, b3, ... bn
  - A series of correlation values is obtained by progressively increasing the size of the lag

# Announcements

### ➤ Papers:

- ➤ Research Proposals due this Wednesday (May 1) no later than 11 pm via D2L
- ➤ Grading Rubric can be seen on D2L
- ➤ Use the stipulated format (check website for details)
- Look at the relevant "guidelines" paper(s) (link on website)
- ➤ Take home final distributed at end of class, due May 6 at 1 p.m. via D2L
- Course Evals now available on D2L
- ≥ 3x5s

### The Basic Idea



See ... CrossCorr\_Sin\_Cos.m

# More Details

- Can be used as a "template matching" procedure
- Compare running average with raw EEG epochs
- This is a method of single-trial signal detection:
- First create a template: either predetermined (e.g., sine wave) or empirically determined (e.g., average)
- Then calculate cross-correlational series between each raw EEG epoch and the template
- If some maximum correlation achieved, conclude signal is present
- If correlation not achieved conclude absent
- This can also be used as a method of determining the latency of a component (by examining the trial-by-trial shifts), or of determining the variability in response for a given individual (again by examining the trial-by-trail shifts)

# Woody's Instantiation

- The Woody Adaptive Filter (Charles Woody, 1967) is a special case and application of cross correlational technique
  The term "adaptive" refers to the fact that the template is not established a priori, but generated and updated by an iterative procedure from the data themselves
- - nitial template is usually either a half cycle of a sine or triangle wave, or the unfiltered average of single trials

    Cross-lagged correlations (or sometimes covariances) are then computed between each trial and this template typically over a limited range of samples (e.g., region of P300, not over "invariant" components)

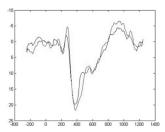
    Each trial is then shifted to align it with the template at the value which yields the maximum cross correlation (or covariance)

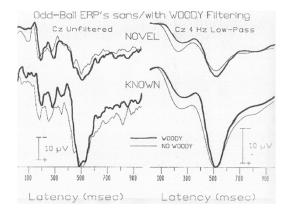
    A new template is then generated by averaging together these time-shifted epochs

    Procedure is repeated using this new average as the template.

  - Procedure is repeated using this new average as the templace are repeated until the maximal values of the cross correlation become stable often, average cross-correlation value increment monitored; if <u>r</u> increases < .005 or .001, then stability achieved
- Some stability active to the s

# Woody Filtering Demo!





# Validity

- Seems to do a fair job of improving signal extraction if a few iterations are used and if the original signal itself is singly peaked
- Wastell(1977) reports a decline in the validity of the procedure if numerous iterations are used
- Therefore, unlike averaging, Woody filtering can only improve signal-to-noise ratio over a definite
- Suggests also that Woody may not be the solution under conditions of very low signal-to-noise ratio

# Principal Components Analysis

- > A method for reducing massive data sets
- ➤ See Handout for gory details

Dimensionality explosions!

32, 64, 128, 256!!!

# PCA (1): The Data matrix

# 

- ➤ Data Matrix above shows only one site could have multiple sites by adding rows for each subject
- > This data matrix is for "temporal PCA" but one could transpose for "spatial PCA"

# PCA (3): The Loading matrix (to guess what components mean)

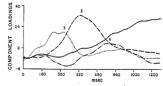
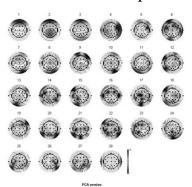


Figure 10-4. Plot of four sets of component loadings derived from a principal-components analysis (PCA) of an ERP data set. Each of the component loading vectors is composed of 128 points corresponding to 128 time points (100-Hz digitizing rate) in the waveforms.

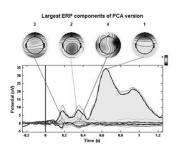
# PCA (2): The Score matrix

- These scores for each subject are optimally weighted composites of the original data, designed to capture as much variance as possible with as few scores as possible.
- ➤ But for conceptual ease, imagine 5 scores: P1, N1, P2, N2, P3 amplitude

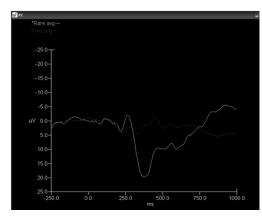
# Spatial PCA on Sample Data



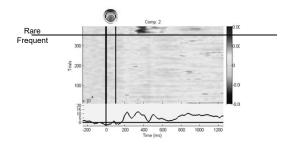
PCA (3b): The Loading Map (for Spatial PCA)



### Reminder: The ERP from which it derives



# PCA Component 2



# PCA (4): Reconstructing Data Matrix

- $\triangleright D_{Nxn} \sim = S_{Nxm} * L_{mxn}$
- ➤ This reconstructed Data matrix will differ slightly from the original Data matrix because not all n components are used.
- To the extent that the m components account for most of the variance in the original data set, the reconstructed data matrix will closely approximate the original data matrix.

# PCA (4): Caveat Emptor

- > PCA is a linear model; assumes the components sum together without interaction to produce the actual waveform
- Sources of variance are orthogonal; if two sources are highly correlated, may result in a composite PCA component reflecting both
- Component invariability in terms of latency jitter across
  - > PCA does not distinguish between variations in amplitude vs variations
  - Especially a problem in comparing control vs pathological groups; pathological groups will typically be more variable
  - Allen & Collins unpublished simulation study:
    - ➤ Two groups: Control & Pathological

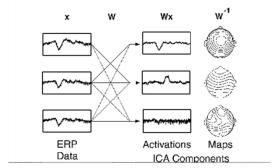
    - I two groups. Coming a randongscar
       I dentical waveforms for each group differed only in latency
       The two groups differed significantly on three of four principal component
    - In other words, if one indiscriminately interprets these as amplitude or morphology differences, one would be WRONG!!!

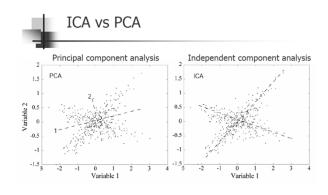
# ICA ... a "better" PCA?

- > PCA finds orthogonal components
  - > First PC accounts for most variance
  - Next PC accounts for most remaining variance
  - > Components will have orthogonal scalp distributions
- ➤ ICA separates temporally independent components
  - Also known as blind source separation
  - ➤ May or may not correspond to brain "hotspots" but do represent functional brain networks
- ➤ See:

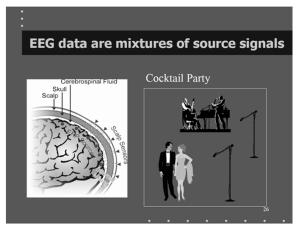
http://arnauddelorme.com/ica for dummies/

# **ICA Decomposition**

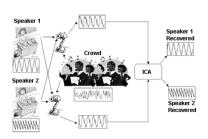


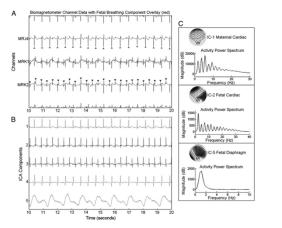


From Tzyy-Ping Jung , presented at EEGLab Workshop, Nov 8,2007

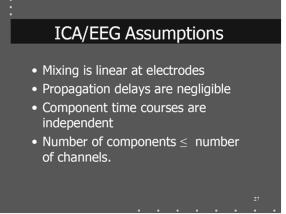


From Tzyy-Ping Jung , presented at EEGLab Workshop, Nov 8,2007



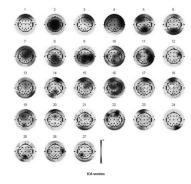


Gustafson, Allen, Yeh, May (2011), Early Human Development

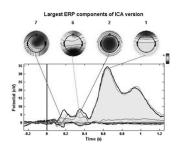


From Tzyy-Ping Jung , presented at EEGLab Workshop, Nov 8,2007

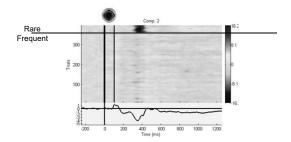
ICA: The Projection Map



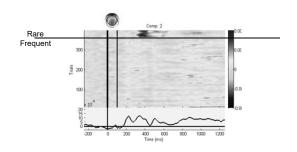
ICA: The Projection Map



ICA: Trial by Trial IC Projection to Pz

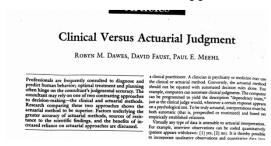


PCA Component 2



ICs as Artifacts!

"Clinical" vs Actuarial Approaches



Dawes, R.M., Faust, D., & Meehl, P.E. (1989). Science, 243, 1668-1674.

# "Clinical" vs Actuarial Approaches

### ➤ Human raters

- ➤ Good source of possible algorithms
- ➤ Lousy at reliably implementing them
  - ➤Inter-rater
  - ➤Intra-rater

### > Actuarial methods

- ➤ Always arrive at the same conclusion
- ➤ Weight variables according to *actual* predictive power

Dawes, R.M., Faust, D., & Meehl, P.E. (1989). Science, 243, 1668-1674.

# ICs as Artifacts!

### ADJUST:

An automatic EEG artifact detector based on the joint use of spatial and temporal features

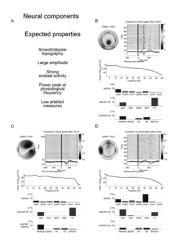
Mognon, Jovicich, Bruzzone, & Buiatti, 2010

# ICs as Artifacts!

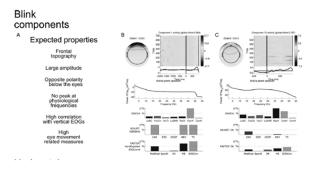
 $MARA \ \ (\text{Multiple Artifact Rejection Algorithm})$ 

able 1 feasures computed by the three automated tools evaluated here. Abbreviations refer to those used in figures and throughout the paper.			
Tool	Artifact type	Measure	Abbreviation
SASICA	Blinks/vertical eye movements	Correlation with vertical EOG electrodes	CorrV
	Horizontal eye movements	Correlation with horizontal EOG electrodes	CorrH
	Muscle	Low autocorrelation of time-course	LoAC or AutoCor
	Bad channel	Focal channel topography	FocCh
	Rare event	Focal trial activity	FocTr
	Non dipolar component	Residual variance	ResVar
	Bad channel	Correlation with Bad channel	CorrCh
FASTER	Eye blinks/saccades	Correlation with EOG electrodes	EOGcorr
	"Pop-Off"	Spatial Kurtosis	SK
	White noise	Slope of the power spectrum	SpecSI
	White noise	Hurst exponent	HE
	White noise	Median slope of time-course	MedGrad
ADJUST	Eye blinks	Temporal Kurtosis	TK
	Eye blinks	Spatial average difference	SAD
	Eye blinks	Spatial variance difference	SVD
	Vertical Eye Movements	Maximum epoch variance	MEV
	Horizontal Eye Movements	Spatial eye difference	SED
	Generic Discontinuities	Generic discontinuity spatial feature	GDSF

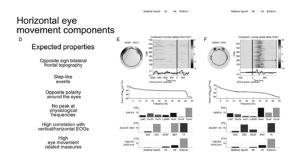
Chaumon et al., 2015



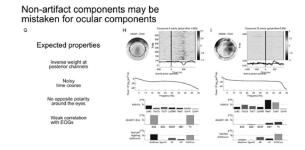
Chaumon et al., 2015



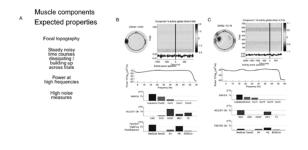
Chaumon et al., 2015



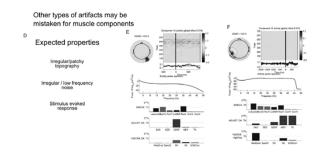
Chaumon et al., 2015



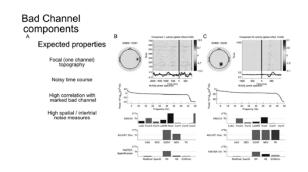
Chaumon et al., 2015



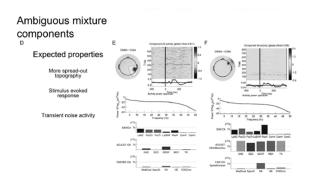
Chaumon et al., 2015



Chaumon et al., 2015



Chaumon et al., 2015



Chaumon et al., 2015

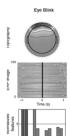
# Rare Events A Expected properties Few high amplitude events in otherwise low amplitude from the properties of the pro

Chaumon et al., 2015

# Eye blinks

- Features used
  - Spatial Average Difference (SAD)
  - Temporal Kurtosis (TK)
- Frontal distribution

High power in delta frequency band

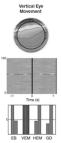


Mognon, Jovicich, Bruzzone, & Buiatti, 2010

# Tyear Sales Same region Look at Component Scroll for what IC 1 looks like Look at Component Scroll for what IC 1 looks like High potentials with these morphology further suggest the IC component is in fact eye blink related

# Vertical Eye Movement

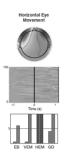
- Features used
  - Spatial Average Difference (SAD)
  - Maximum Epoch Variance (MEV)
- Frontal distribution similar to that of an eye blink



Mognon, Jovicich, Bruzzone, & Buiatti, 2010

# Horizontal Eye Movement

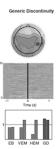
- Features used
  - Spatial Eye Difference (SED)
  - Maximum Epoch Variance (MEV)
- Frontal distribution in anti-phase (one positive and one negative)



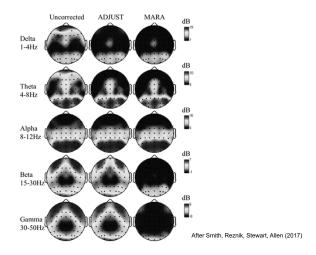
Mognon, Jovicich, Bruzzone, & Buiatti, 2010

# Generic Discontinuities

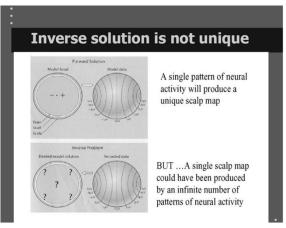
- Features used
  - Generic Discontinuities Spatial Feature (GDSF)
  - Maximum Epoch Variance (MEV)
- Variable distribution
- Sudden amplitude fluctuations with no spatial preference
  - Could be present in as little as one or 2 trials, and limited to 1 channel
- In component data scroll weird activity in the trial plotted on the IC activity







# Neural Sources of EEG



From Tzyy-Ping Jung , presented at EEGLab Workshop, Nov 8,2007

# Source Analysis

- ➤ BESA -- Brain Electrical Source Analysis
- ➤ This is a model-fitting procedure for estimating intracranial sources underlying ERPs
  - ➤ Estimate -- if model fits, then data are consistent with these sources; yet there is no unique solution
  - ➤ Not for ongoing EEG -- too many sources

### **BESA**

➤ Imagine a data matrix of ERPs:

 $V_{Cxn}$  (# Channels by # timepoints)

Note that this is really the result of the subtraction of the activity at the reference from the activity at the these sites; i.e.,

 $\mathbf{V}_{\mathbf{Cxn}} = \mathbf{U}_{\mathbf{Cxn}} - \mathbf{R}_{\mathbf{Cxn}}$ 

Note: the reference matrix has identical rows! Thus BESA Presumes that all channels referenced to the same reference!

# **BESA**

> Reconstruct a data matrix that includes not only the original channels, but the implicit channel (reference) as well:

 $U_{Exn}$  (# electrodes = # channels+1),

which represents the activity at each electrode with respect to an average reference (i.e., the average of all channels)

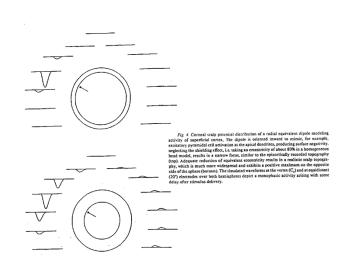
# **BESA**

- $\triangleright$  Now this matrix  $U_{Exn}$  can be decomposed
  - $\triangleright$  a set of sources:  $S_{Sxn}$  (# Sources by # timepoints)
  - $\triangleright$  a set of attenuation coefficients  $C_{ExS}$
  - ightharpoonup so that  $U_{Exn} = C_{ExS} S_{Sxn}$

# **BESA**

- > The attenuation matrix is determined by:
  - > the geometry between the source and the electrodes
  - > the nature of the conductance of the three-layer head model (Brain, Skull, Scalp);
    - > the skull is less conductive than the layers on either side
    - > this results in a spatial smearing of potentials as they cross the skull
    - the skull produces the equivalent of a brain that is 60% of the radius of the outer scalp (rather than the "true" figure of ~84%)

Next

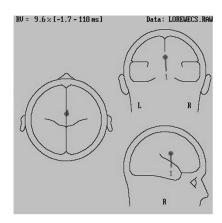






# **BESA**

- Note that the decomposition of U into C and S
  - ➤ an electroanatomical time-independent matrix (C) that reflects that anatomical substrates do not move around in the head
  - ➤a time-variant dipole source potential matrix that represents the change in activity of each source over time



# BESA Vs PCA Vs ICA (a battle of acronyms)

- ➤ This decomposition is akin to PCA/ICA
  - > PCA and ICA have sources and propagation coefficients
  - PCA solutions are constrained by orthogonality of components, and by those that account for greatest common variance
  - ➤ ICA constrained to find temporally independent components
  - ➤ BESA solutions are constrained by the geometry of the head, the volume conduction of the dipoles, and the anatomical constraints dictated by the user (e.g., inside the head, symmetrical, not in the ventricles, must not be in the brainstem after a certain point in time, etc...)

### BESA Vs PCA Vs ICA continued

- ➤ Like PCA/ICA, the reconstruction of the original data set will be imperfect
  - ➤ With all methods. better chance of reconstructing the original matrix if data are reliable
  - > If you capture the important sources, the reconstruction should be very good (i.e., small residual variance)
  - > It is useful to attempt to upset a solution by inserting another source and seeing if:
    - $\triangleright$  the original solution is stable
    - > the new source accounts for any substantial variance
- Can do dipole localization (BESA) on an IC!

# Dipole Fitting PCA ICA ICA

# You can try it!



# **Implementations**

- BESA can be used:
  - > in a strict hypothesis-testing manner by designating sources a priori and testing the fit
  - in an exploratory/optimizing manner by allowing the program to iteratively minimize the residual variance (between observed and reconstructed waveforms) by:
    - > moving dipoles
    - > changing the orientation of dipoles
    - > altering the time-by-activity function of the dipoles

# BESA – Did it work?

- ➤ In the end, the adequacy of your solution will be judged by
  - > stability of your solution:
    - > against insertion of additional dipoles
    - > across multiple subjects
  - > anatomical feasibility
  - > follow-up tests with patients with lesions
  - your reviewers!

# Recording EEG in fMRI environments: Oodles of Issues

- ➤ EEG can be bad for fMRI
  - ➤ Wires and electrodes can be ferromagnetic = TROUBLE
  - ➤ Wires and electrodes can be paramagnetic = less trouble
- > MRI and fMRI can be bad for EEG
  - > Gradient switching creates huge artifact for EEG
  - Movement in Magnetic fields creates current in any conductive medium (e.g. wires!)
  - ➤ High frequency current can make wires HOT and RF is 127.68 MHz at 3T that's fast, and can create mega-hurts!
    - ➤ Thus in-line 10K resistor

# Special Caps



- ➤ Need conductive material
- > That will not heat up
- ➤ That will not pose hazard in strong magnetic field
- ➤ That includes inline resistor to prevent any induced current from reaching the subject
- ➤ That includes Styrofoam head at no charge

# Whence EEG Artifacts in fMRI?







2. Protons wobble in alignment with magnetic fields of varying intensity; frequency of wobble is proportionate to strength of individual magnetic field



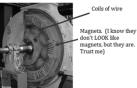
A brief radio signal, nose soundwave frequency quals the frequency of obble of certain protons, nocks those protons out

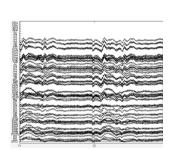


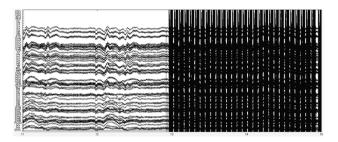
4. When radio signal cease protons snap back into alignment with magnetic field, emitting a radio sign of their own, that announce the presence of a specific

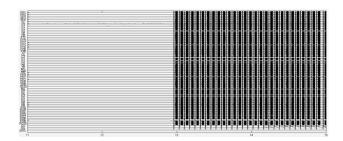
# Whence EEG Artifacts in fMRI?

- → Faraday's law of induction...
  - → induced electromotive force is proportional to the time derivative of the magnetic flux
  - → Flux = summation of the magnetic field perpendicular to the circuit plane over the area circuit
  - $\bullet$   $\varepsilon = d\Phi/dt$
- → Can reflect:
  - → changes in the field
  - Changes in the circui relative to the field







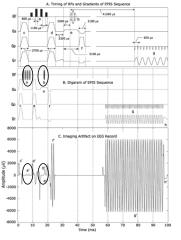


# Whence EEG Artifacts in fMRI?

- → RF pulses
  - → For 3T = 127.6 MHz
  - → Brain oscillations ≈ 0.5-50 Hz
  - → Amplifier frequency range = DC-3.0 KHz
- → Artifacts thus attenuated, but still range overwhelm the EEG signal

# Whence EEG Artifacts in fMRI?

- → Gradient Switching
  - → Artifact approximates differential waveform of the gradient pulse
  - → Polarity and amplitude varies across channels
  - → Frequency ≈ 500-900 Hz
- → EEG dominated by
  - → harmonics of slice repetition frequency (≈10-25 Hz)
  - → convolved with harmonics of volume repetition frequency (≈0.2-2 Hz)
- → Artifacts in range from 1000-10,000 μV!



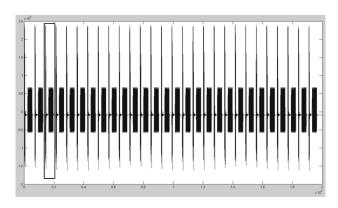
- = radiofrequency wave;

- Gs = slice selection gradient
  Gp = phase encoding gradient
  Gr = readout gradient
  a = Fat suppression pulses (1-3-3-1 pulses)
  b = slice selection RF

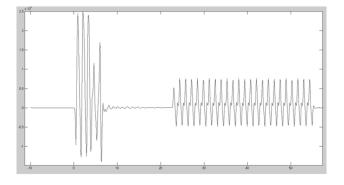
- b = slice selection RF c, d, h = spoilers e = slice selection gradient f = dephasing and rephasing gradient g = readout gradient ' = EEG artifact corresponding to letter

Ritter et al., 2009

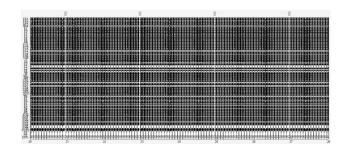
# Average Artifact (across 1 TR)



# Average Artifact (0-60 msec)

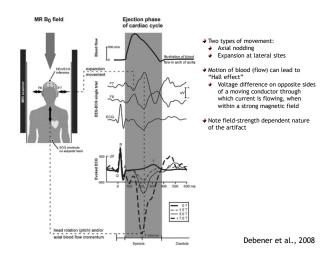


# Artifact (across several TRs)

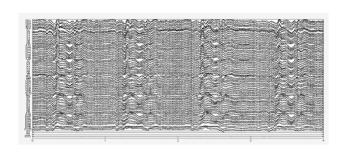


# Whence EEG Artifacts in fMRI?

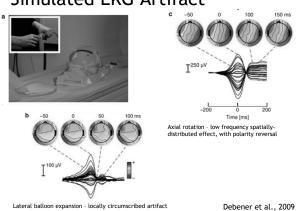
- → Faraday's law of induction...
  - → induced electromotive force is proportional to the time derivative of the magnetic flux
  - Flux = summation of the magnetic field perpendicular to the circuit plane over the area circuit
  - $\bullet$   $\varepsilon = d\Phi/dt$
- → Can reflect:
  - → changes in the field (gradient switching, RF)
  - → Changes in the circuit geometry or position relative to the field due to body motion



# EEG in Magnet (no scanning)



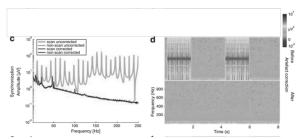
# Simulated EKG Artifact



Ohmagawd... Help me in

# **REMOVING THOSE PESKY ARTIFACTS!**

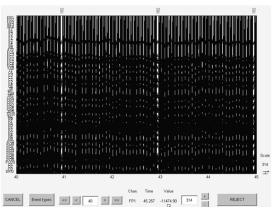
# Gradiant/RF removal via moving average subtraction



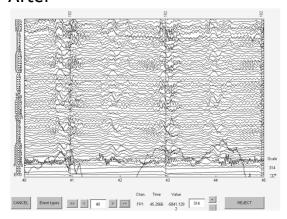
# FASTR: FMRI Artifact Slice Template Removal

- → Part of FMRIB Plug-in for EEGLAB
- → Upsample to at least 20K Hz
- → Align slices for slight jitter in timing
- → Moving Window approach with subtraction
- → PCA on artifact residuals form Optimum Basis Set (OBS) to reduce residual artifacts by 90%
- → Downsample to original rate
- → Sample Results.....

# **Before**



# After



# Alternatively ... BrainVision

- → Sync EEG clock to MR clock
- → No jitter in timing, no need to upsample (recorded at 5000 hz)
- → Moving Window approach with subtraction
- → Downsample to original rate
- → Sample Results......



# ECG-related removal via moving average subtraction (Allen et al. 1998)

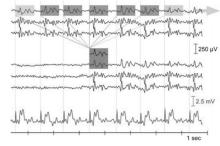
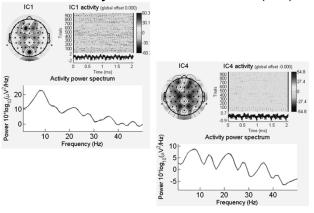
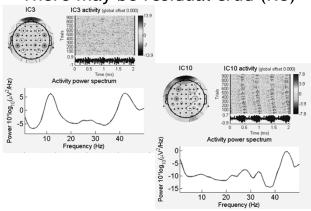


Fig. 5 Schematic of the average artefact subtraction procedure. For each channel, a waveform template is generated by averaging EEG opechs over adjacent cardiac eyeles, with the time locking event being derived from the ECG. The template generation is combined with a moving average procedure, and new templates are generated for each cardiac cycle. The procedure is repeated for each EEG channel

# There may be residual crud (RC)



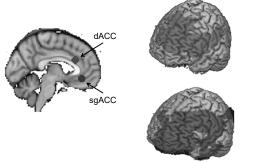
# There may be residual crud (RC)



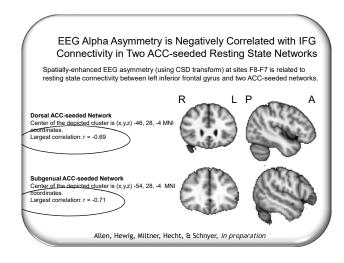
Simultaneous EEG and RSfMRI (following ICA!)

# Multi-modal Imaging

→ Create RS-fMRI network with ACC seeds



Allen, Hewig, Miltner, Hecht, & Schnyer, in preparation



# **EEG-fMRI** Synopsis

- → Less relative left frontal activity (indexed by EEG) is related to increased connectivity of left IFG to two ACC-seeded RS networks
- → Consistent with:
  - → Hyper-connectivity in RSfMRI emotion networks in MDD (e.g., Grecius et al., 2007; Sheline et al., 2010)
  - → Frontal EEG asymmetry findings of less relative left frontal activity in risk for MDD.
- Alpha power may regulate network connectivity
  - → Note: Between vs Within Subjects

# Within Subjects' Moderation of RSfMRI Connectivity

- → Calculate F8-F7 alpha asymmetry for each TR
  - ◆ EEG leads TR by 4.096 seconds
- → Median split into high (left) and low (right)
- → Entered as moderator in PPI approach (cf. Friston et al., 1997)
  - Tests whether strength of connectivity to seed region varies as a function of the moderator

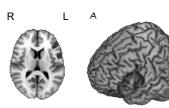
Allen, Hewig, Miltner, Hecht, & Schnyer, in preparation

# BETWEEN-SUBJECTS' DATA DOES NOT NECESSARILY SUPPORT A WITHIN-SUBJECTS' INTERPRETATION

# Within Subjects' Moderation of RSfMRI Connectivity



Dorsal ACC Seed



Greater Connectivity with Less Left Frontal Alpha or Greater Left Frontal Alpha

Allen, Hewig, Miltner, Hecht, & Schnyer, in preparation

# Within (red) and Between (blue) Within-subject effects more extensive

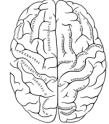


# Cognitive Control over Emotion

 IFG has a key role in mediating the success of cognitive control over emotional stimuli

# Cognitive Control over Emotion

→ Left IFG: Language and self-referential processing



- Right IFG: Attentional control
- → behavioral inhibition
  - → suppression of unwanted thoughts
  - → attention shifting
- efforts to reappraise emotional stimuli
- → Working Hypothesis:
  - Hyperconnected left IFG and emotion networks: rumination
  - Hypoconnected right IFG: difficulty disengaging from emotion



# Psychophysiology -- Synopsis

- ➤ Psychophysiology is inherently interdisciplinary, and systemic
- ➤ Principles learned here can apply to a wide range of physiological signals
  - **≻**Recording
  - **≻**Processing
  - **▶**Interpretation

# Psychophysiology -- Synopsis

- Ultimately we obtain correlates of behavior and experience
  - ➤ Psychophysiological Correlates are not privileged; they are no better, no worse, than any other correlate of behavior and experience
- ➤ The utility of these correlates like any correlates in science hinges upon:
  - > good experimental design
  - > strong theoretically driven hypothesis testing
  - > the development of a nomological net, a set of interrelationships among tangible measures and constructs that place the findings in a larger theoretical context, and lend construct validity to the measures and findings