

Advanced Signal Processing I

Digital Filters

Time Frequency Approaches

Ocular Artifacts

Announcements

- Change in due date:
 - Research Proposals due next 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 next week, due May 6 at 1 p.m. via D2L
- Course Evals now available on D2L
- Today:
 - No office hour (but could be available at 5)
- 3x5s x 2

Advanced Signal Processing I

Digital Filters

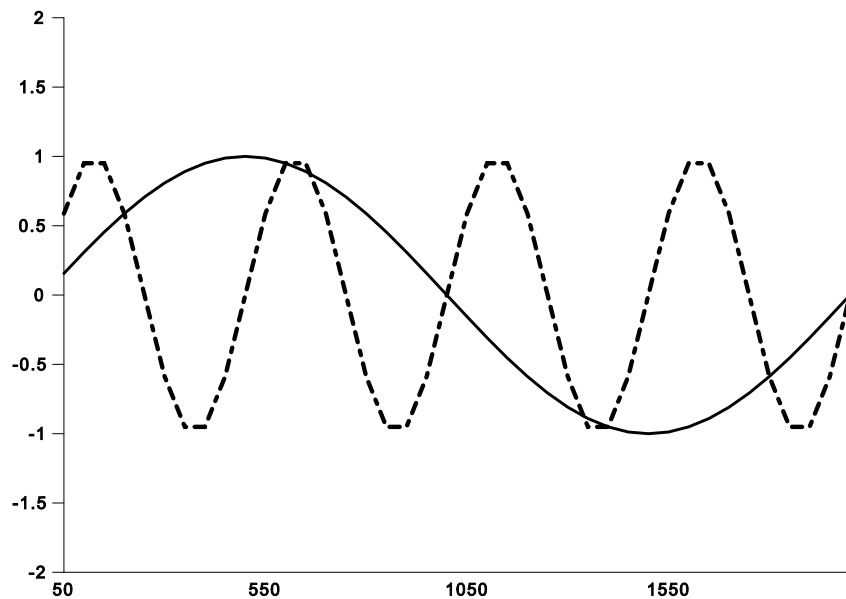
Time Frequency Approaches

Ocular Artifacts

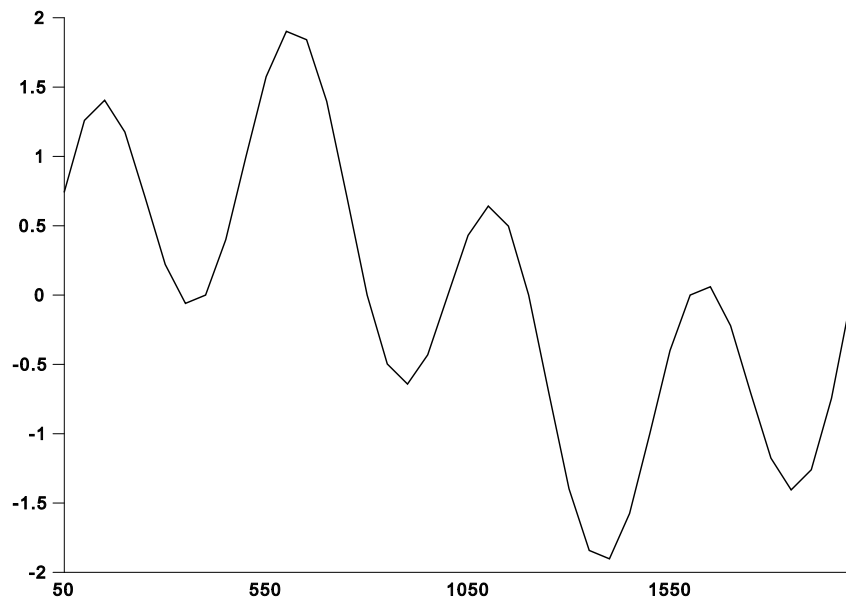
Digital Vs. Analog Filtering

- Analog filters can introduce phase shift or lag
 - Certain frequency components "lagging" behind the others
 - This is the effect of a capacitor literally slowing a signal
 - Some frequencies are slowed more than others
 - Problem: some ERP components could be distorted
- Analog filters are irreversible – once applied, there's no turning back
- Hence, digital filtering is a preferred alternative.
 - No phase shift
 - Is widely used in last several decades
- If digitized signal has minimal filtering, nearly infinite possibilities exist for digital filtering later

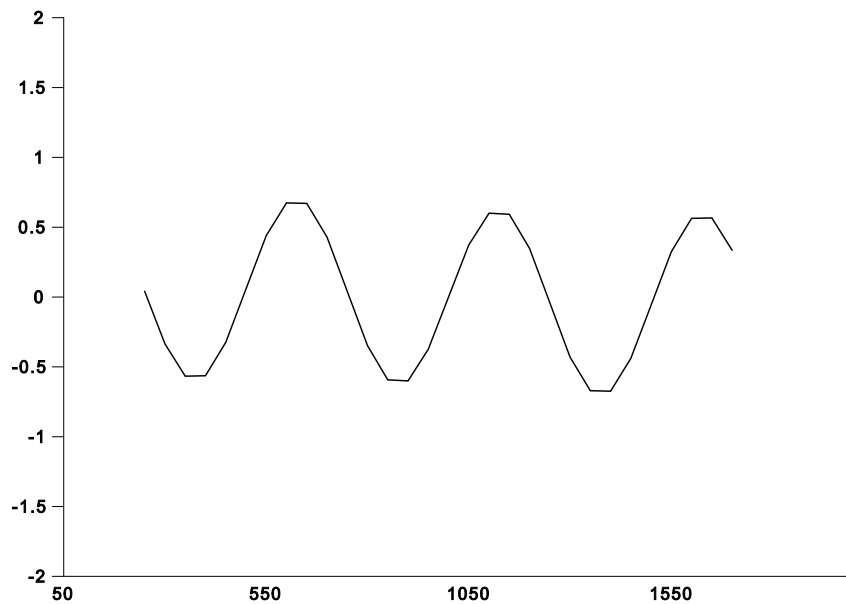
Constituent Waveforms



Resultant Waveform

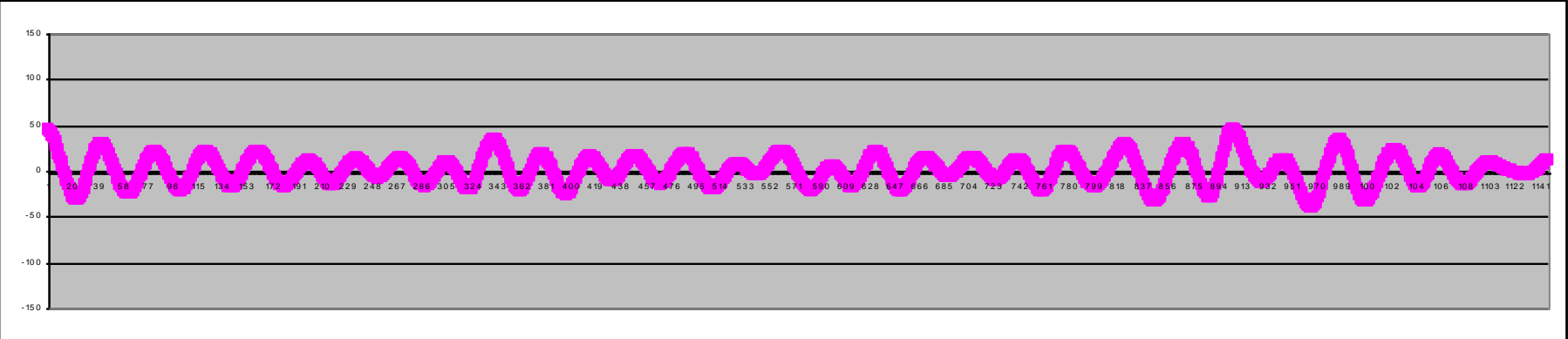
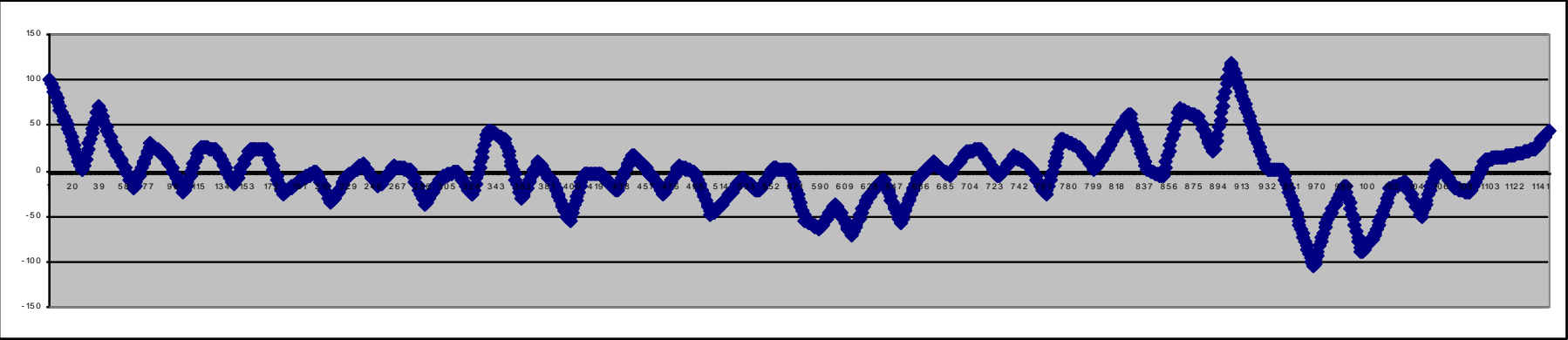


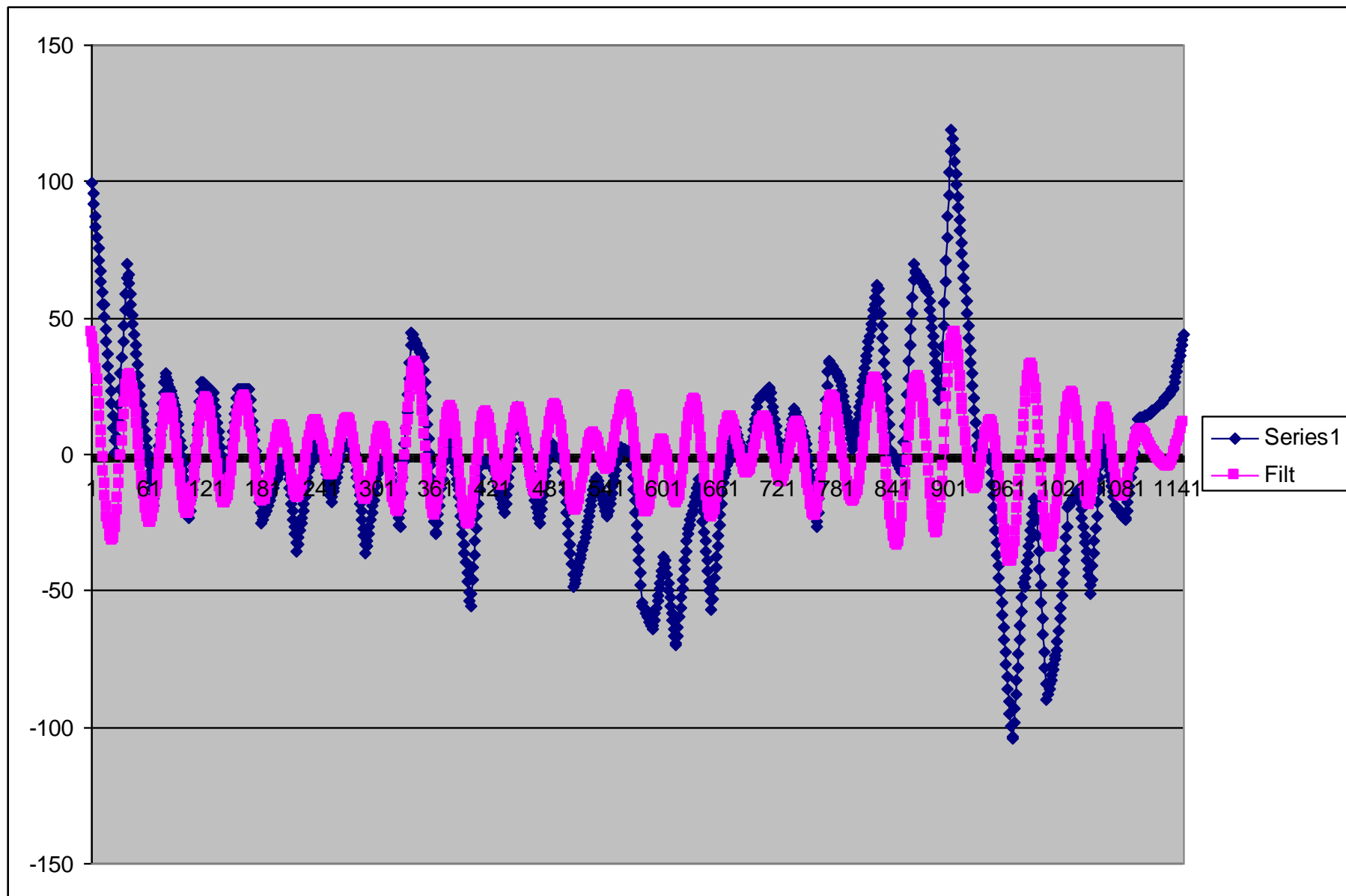
High Pass Filtered



Low Pass Filtered







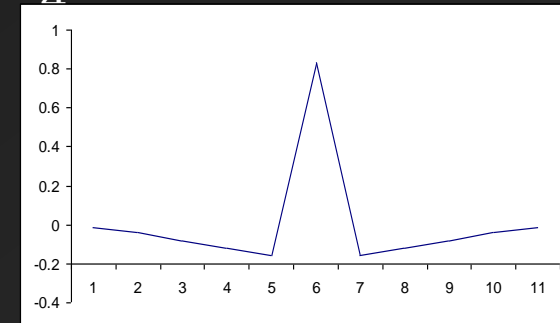
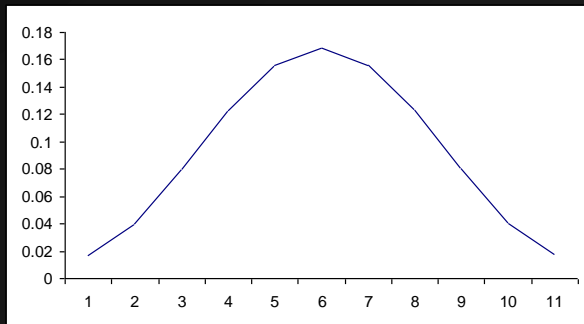
The Details!

- Handout on Digital Filtering

Filter Details

A. Linear digital filters may be conceived of as vectors of weights that are to be multiplied by the digitally sampled values from a waveform. The filters given below are both 11 point digital filters with a half-amplitude frequency cutoff of approximately 17.5 Hz for data sampled at 200 Hz.

LOW PASS		HIGH PASS	
COEFFICIENT	LAG	COEFFICIENT	LAG
-----	---	-----	--
0.0166	5	-0.0166	5
0.0402	4	-0.0402	4
0.0799	3	-0.0799	
0.1231	2	-0.1231	
0.1561	1	-0.1561	
0.1684	0	0.8316	
0.1561	-1	-0.1561	
0.1231	-2	-0.1231	-2
0.0799	-3	-0.0799	-3
0.0402	-4	-0.0402	-4
0.0166	-5	-0.0166	-5



More Details

- 11 point filters indicates that 11 sample points are used in the determination of the new filtered value of any one sample point
- Middle (sixth) sample point is a weighted sum of the first 11 samples.
- The non-recursive filter uses raw sample values in the calculations; recursive filters use the already filtered values of preceding samples in the calculations. Non-recursive filters are more straightforward and more commonly used.
- The term linear denotes that the filter involves the computation of weighted sums of the digital sample values. Other filtering algorithms can be devised, but are less often applied to psychophysiological signals.

More Details (cont')

- Digital filters have characteristics that are sampling-rate dependent.
- These same filters would have a different cutoff frequency for data sampled at different sampling rates.
- Once you know the characteristics of a digital filter at a given frequency, it is a simple matter to convert the filter to another sampling rate as follows:

$$17.5/200 = x/1000 ; x = 87.5 \text{ @ } 1000 \text{ Hz Sampling rate}$$

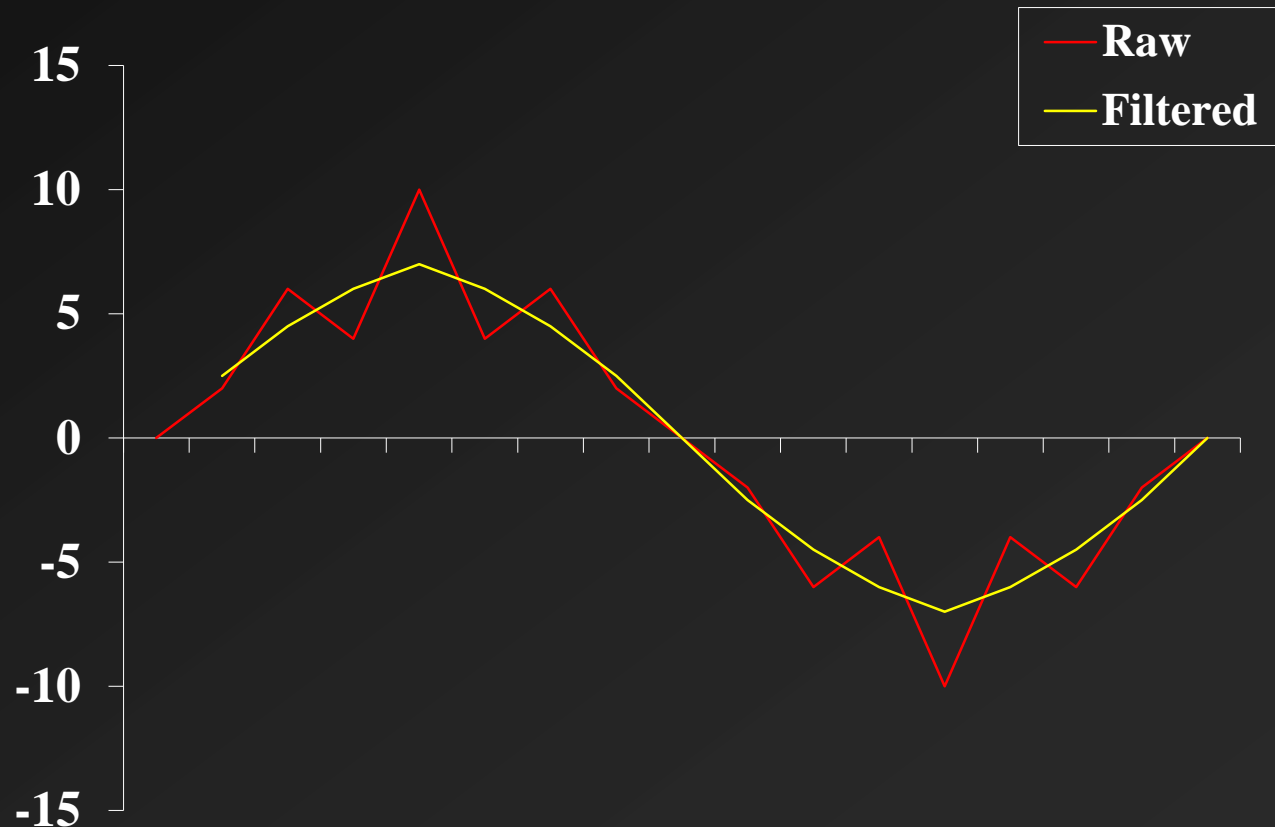
$$17.5/200 = x/20 ; x = 1.75 \text{ @ } 20 \text{ Hz Sampling rate}$$

Muy Simple Filter

[.25 .5 .25]

To apply: Iterate through data segments the size of the filter

$$\text{filt}_{1 \times 3} * \text{segment}_{3 \times 1} = \text{filteredpoint (scalar)}$$



Some filters and their Transfer Functions

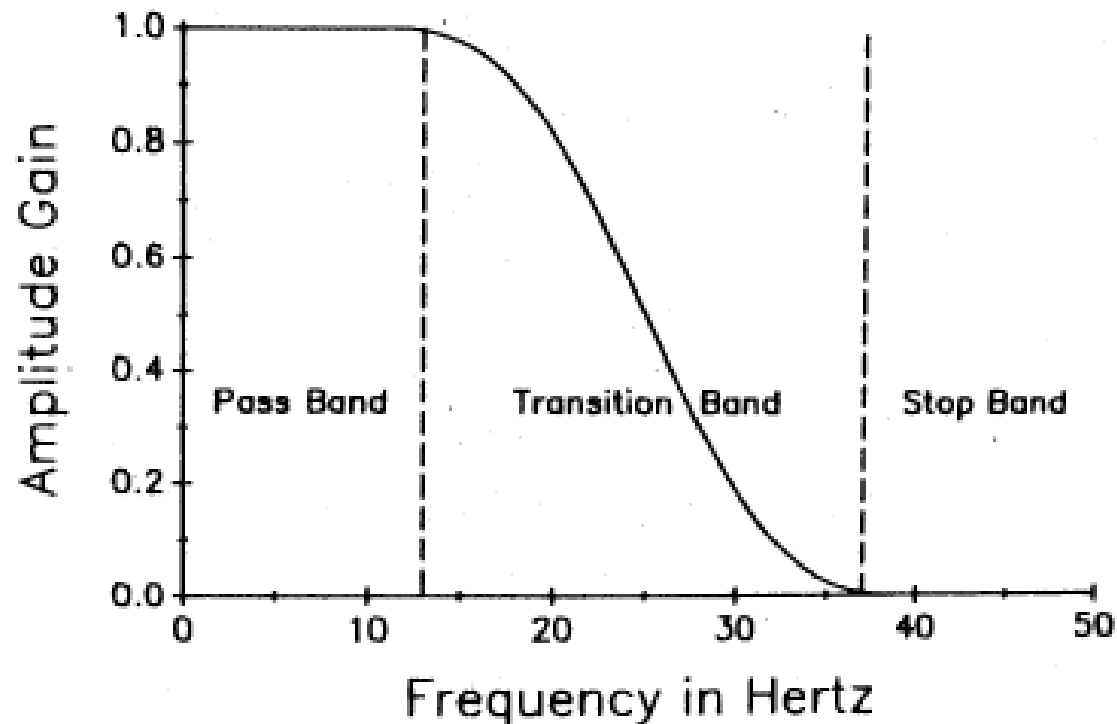
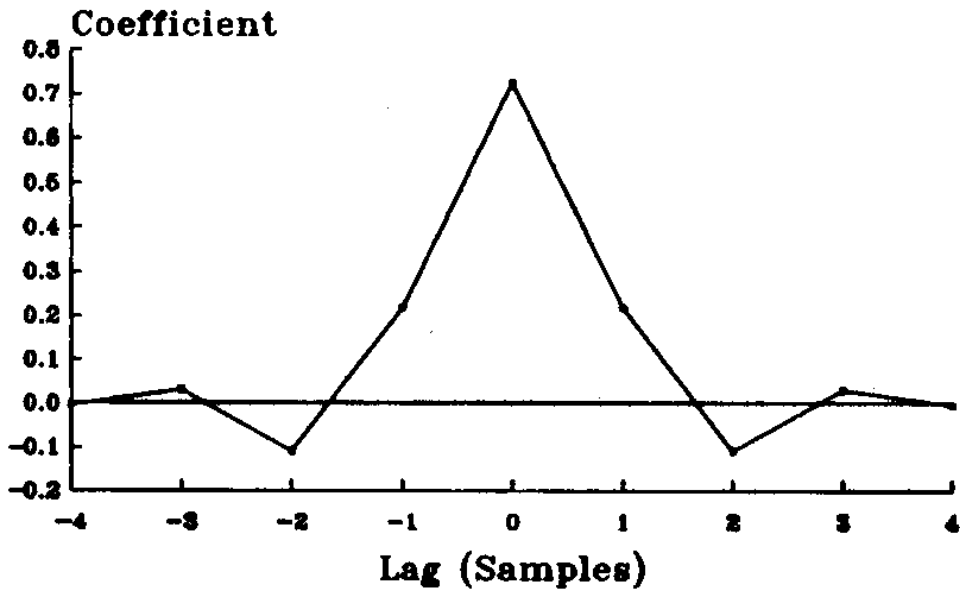


Figure 1. The gain function of a filter is divided into the pass band, transition band, and stop band. The gain function shown is for a low-pass filter.

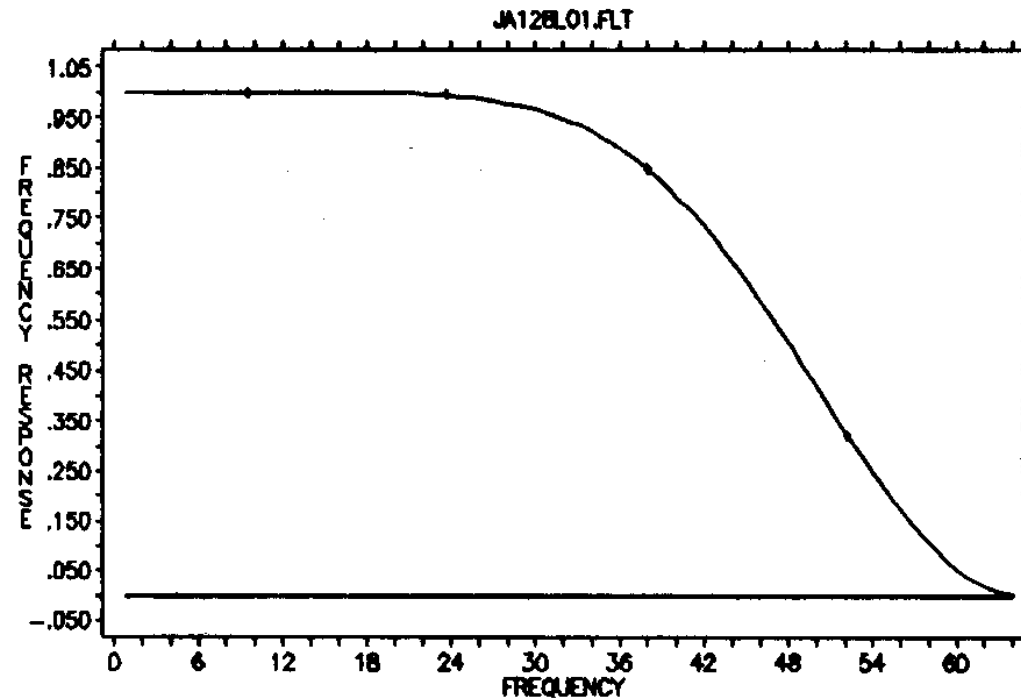


Impulse Response

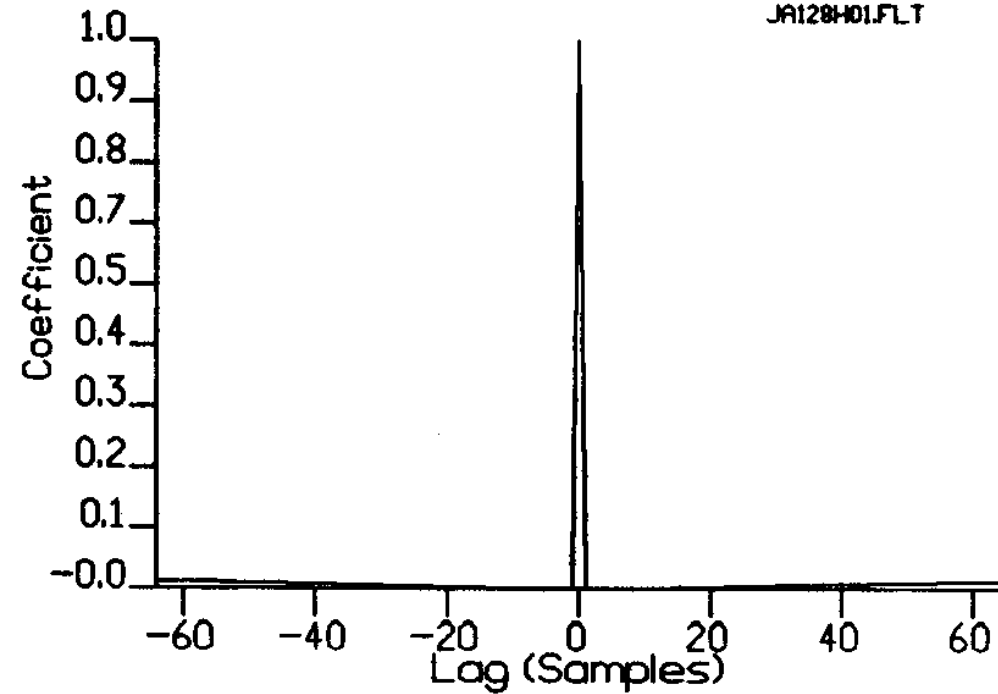
Note:

- FFT of Impulse Response (filter) gives transfer function
- Inverse FFT of transfer function yields impulse response (filter coefficients)

Transfer Function

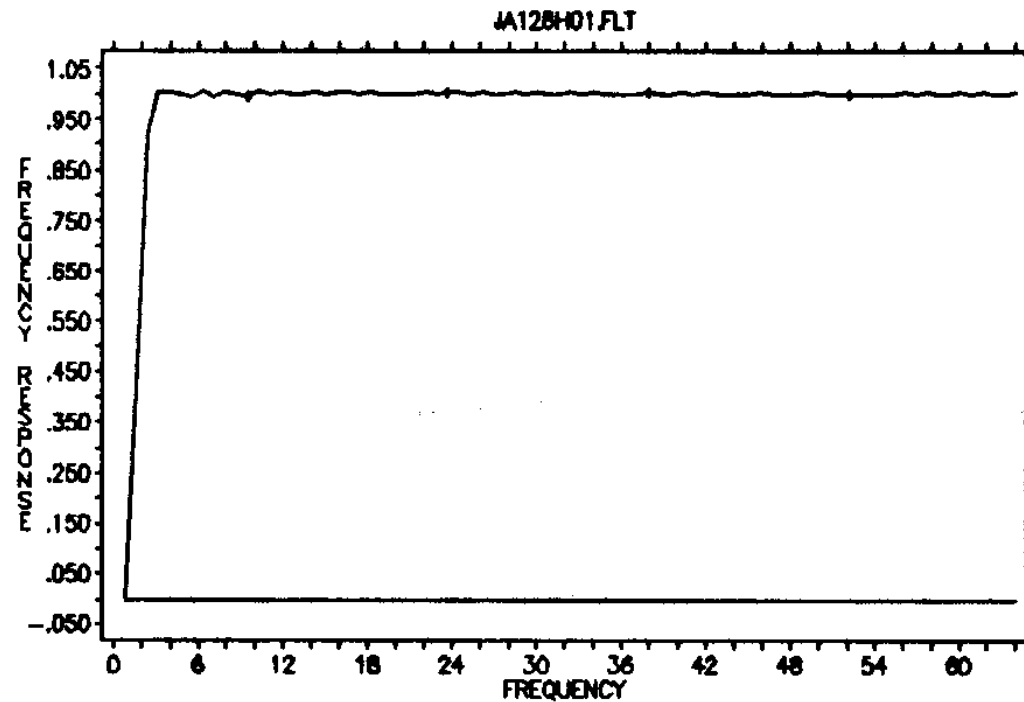


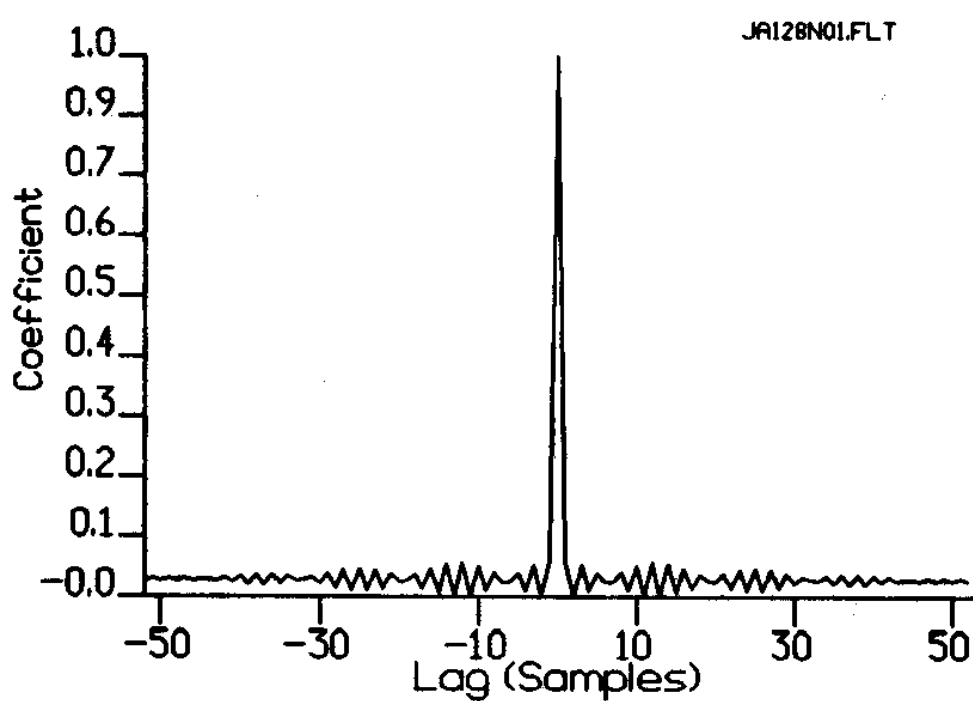
JA128H01.FLT



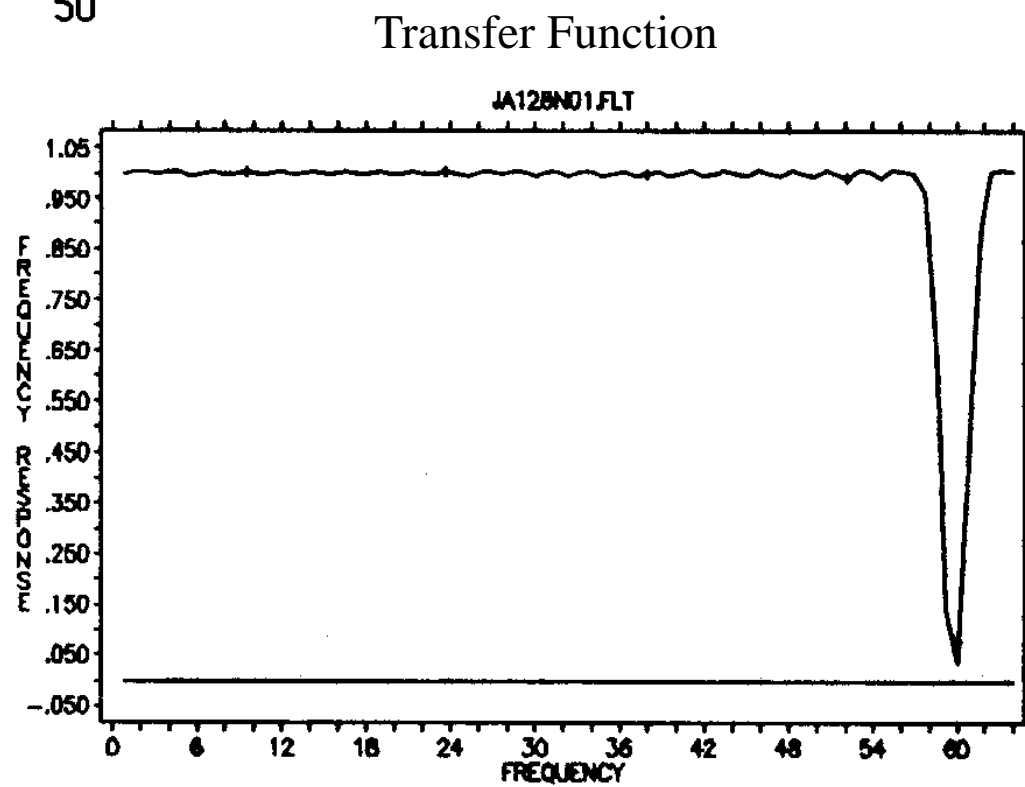
Impulse Response

Transfer Function

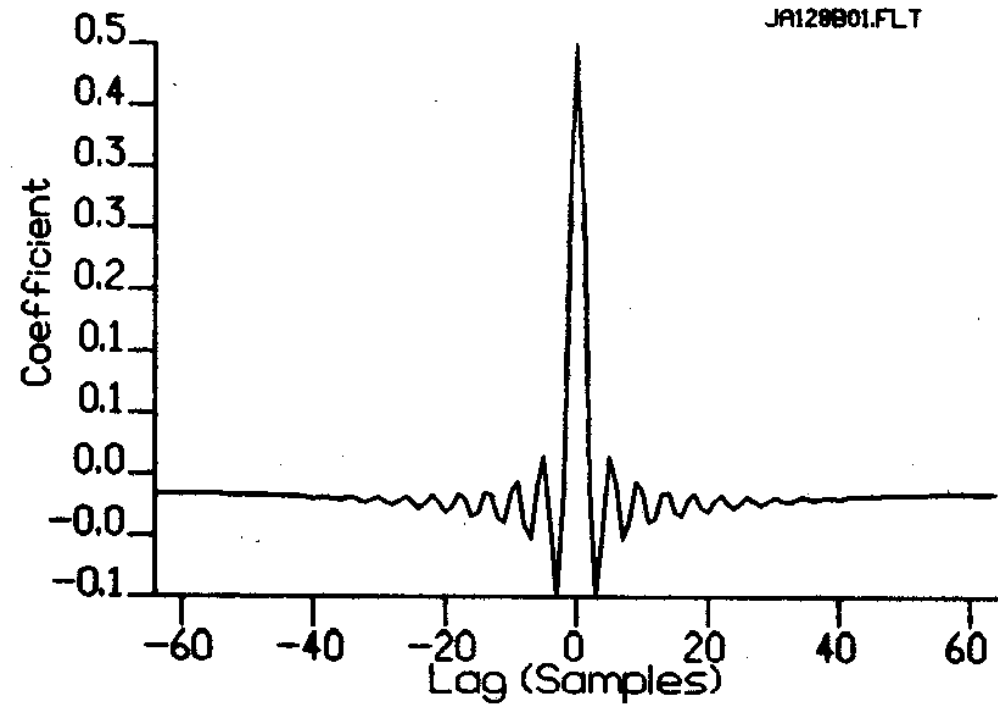




Impulse Response



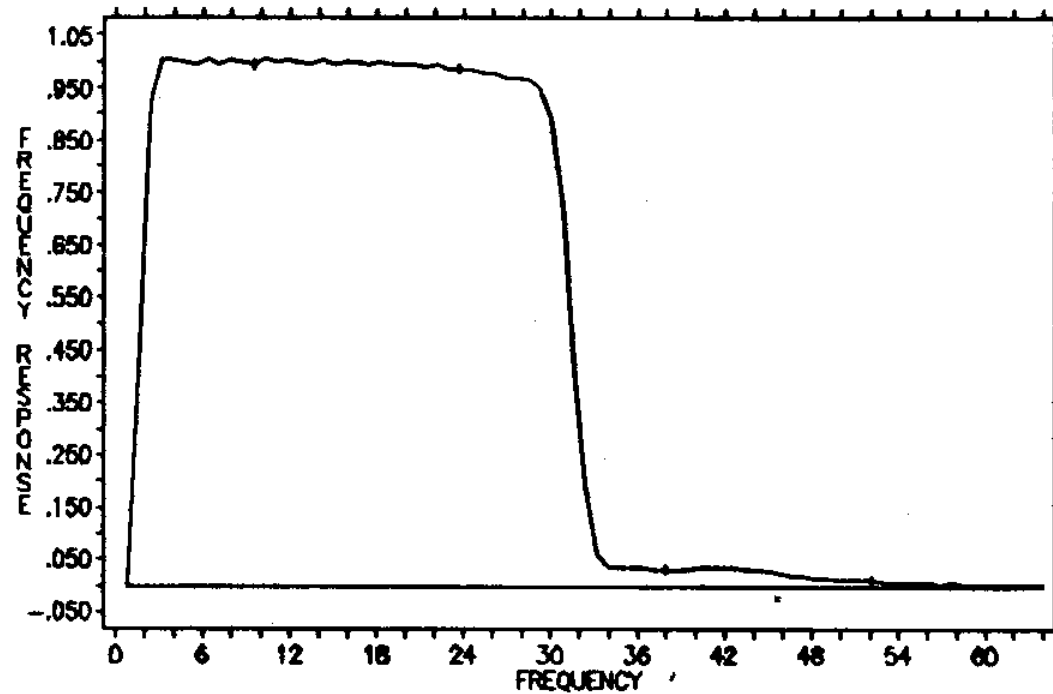
JA128B01.FLT



Impulse Response

Transfer Function

JA128B01.FLT



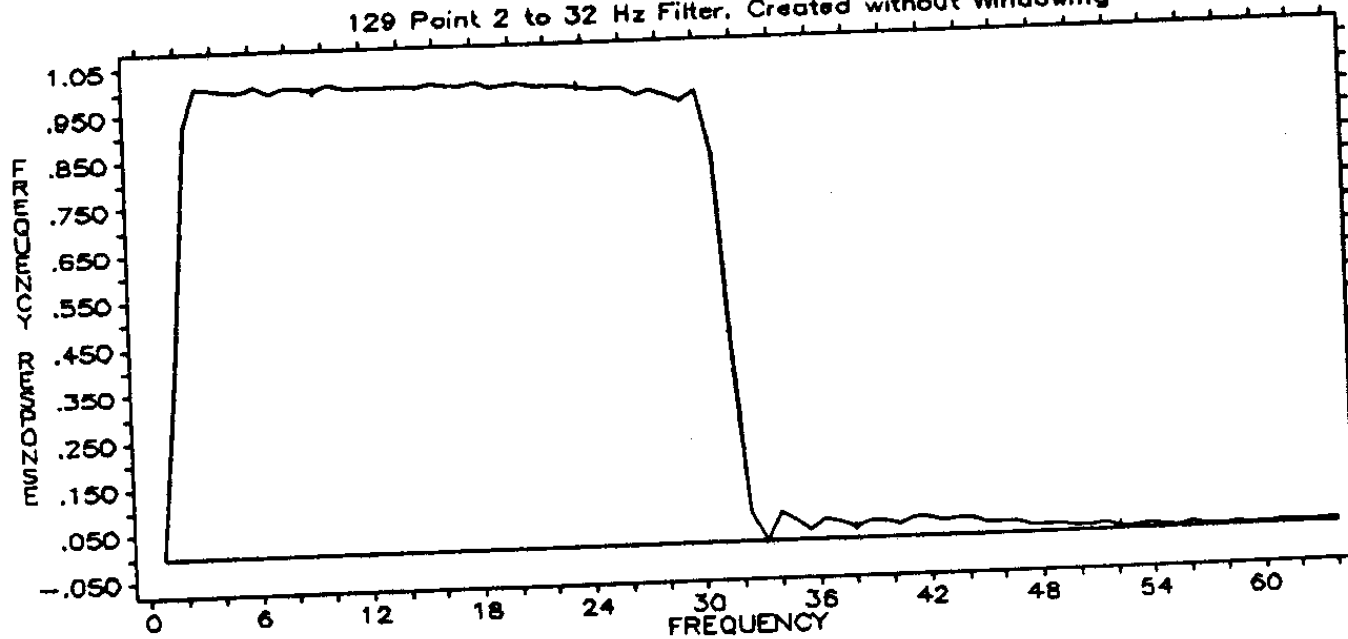
Pragmatic concerns

- Sample extra data points; many if you want sharp roll-off
 - The filter cannot filter the first $(n-1)/2$ points for filter length n
- Try out your filter via FFT analysis or via derivation of the transfer function before you apply it routinely

Convolution of Filters

- If you have filters that do desirable things, but neither does it all, you can convolve filters upon one another
- Since filter's have endpoints near 0, you can "pad" the ends with 0's so as not to lose data points
- Windowing an option

129 Point 2 to 32 Hz Filter. Created without Windowing



The effects of windowing on broadening the transfer function, but reducing bandpass ripple

Hamming Taper, for i coefficients $-j$ to $+j$,

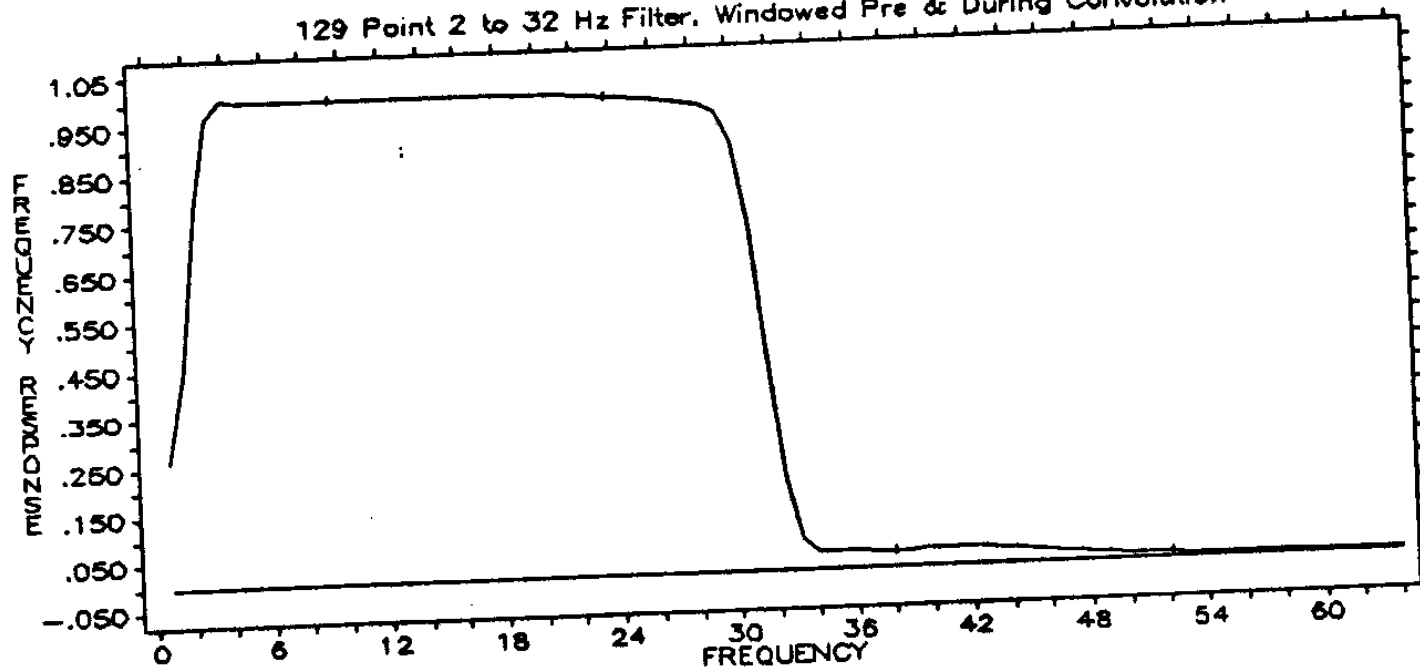
$$\text{WinFilt}(i) = \text{NonWinFilt}(i) * w_i$$

where:

$$w_i = .54 + .46 * \cos(\pi p_i)$$

$$p_i = i/(j+1)$$

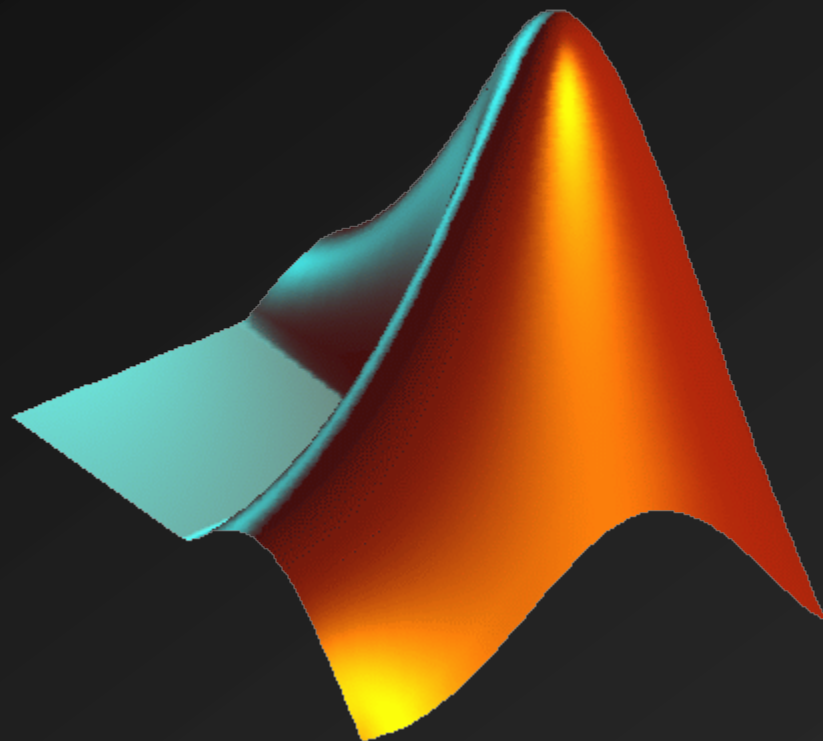
129 Point 2 to 32 Hz Filter. Windowed Pre & During Convolution



Use in Single Trial Analysis

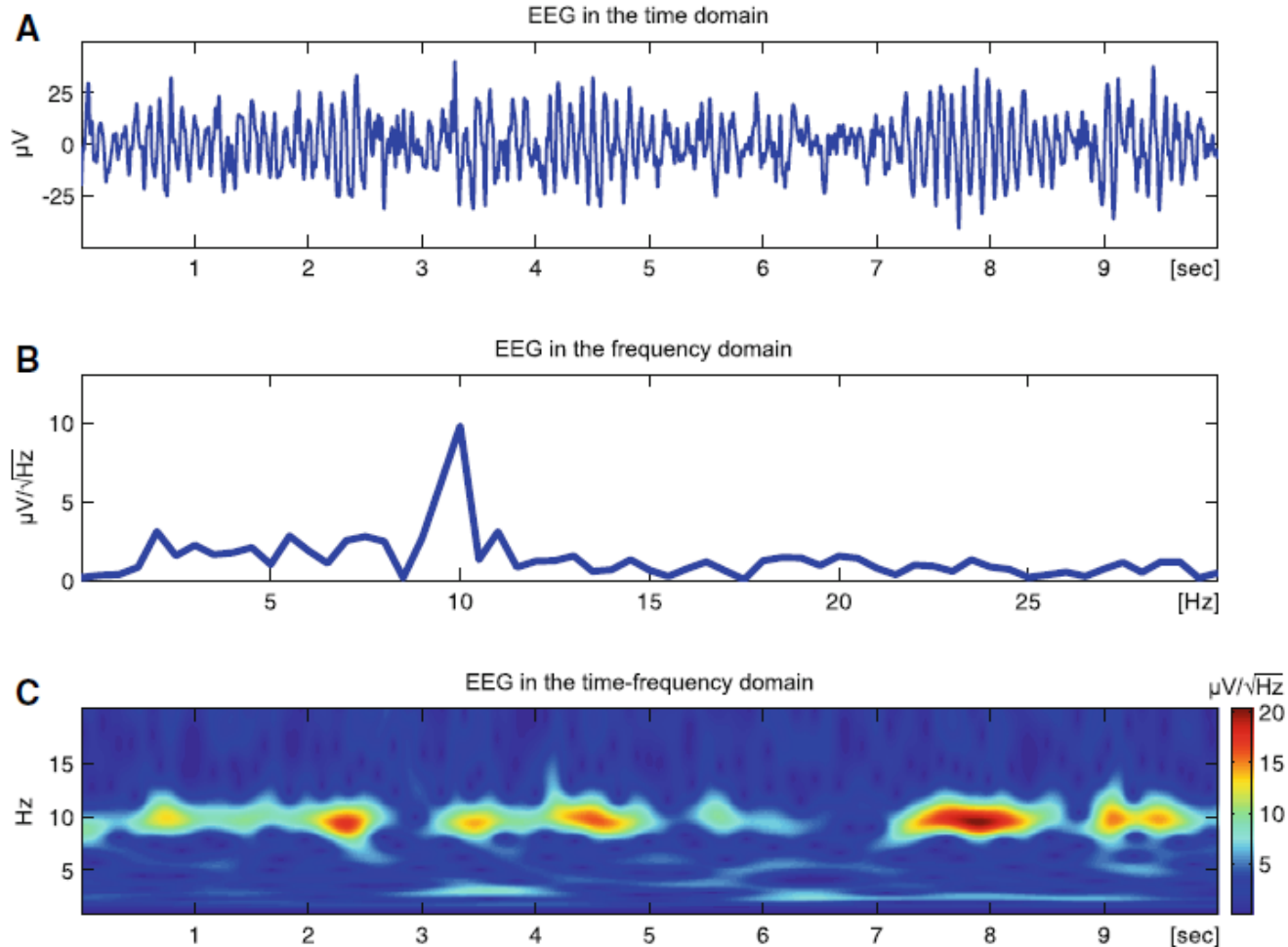
- With stringent digital filtering, you may be able to discern peaks on an individual trial basis

Digital Filtering and More!



Time-Frequency Approaches

Brain Topogr (2014) 27:438–450

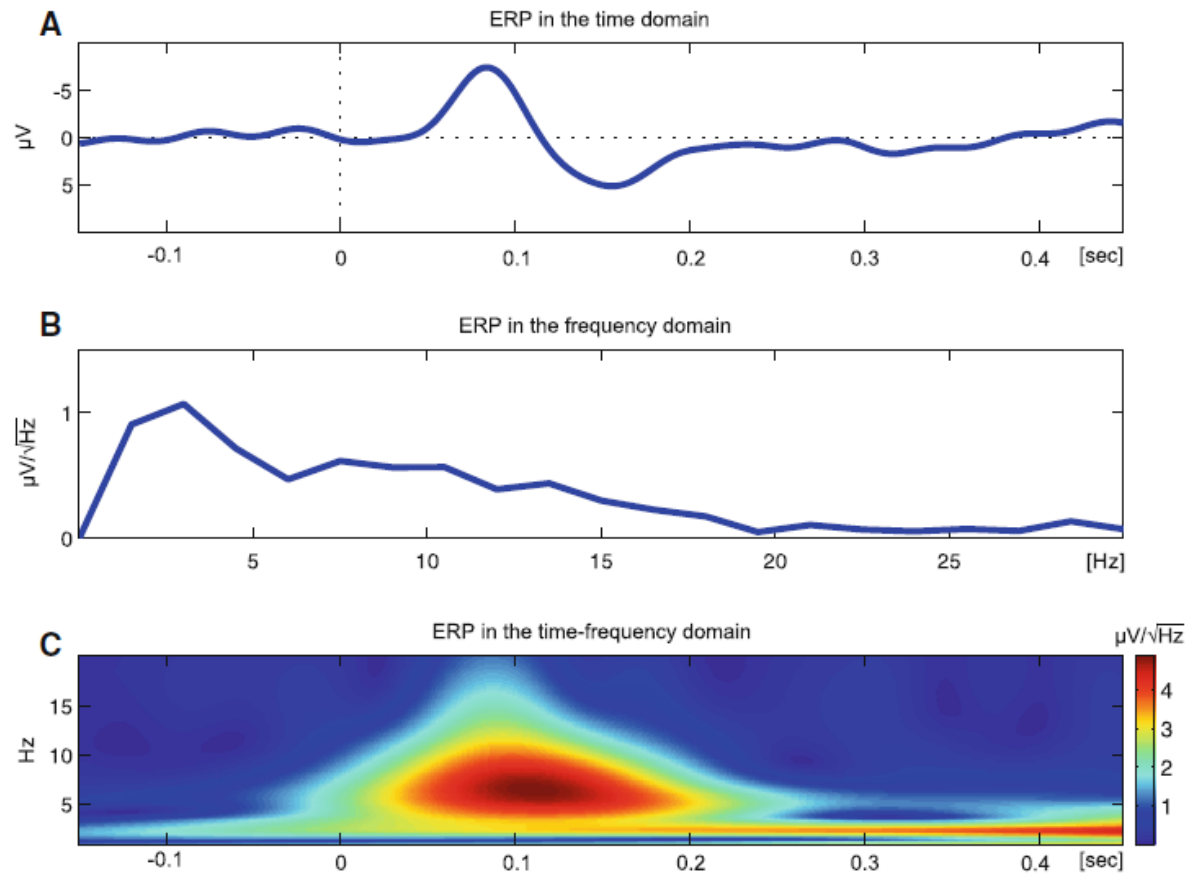


Let's make sure we understand Time-Frequency Space!

[MUSICLAB.CHROMEEXPERIMENTS.COM/SPECTROGRAM](https://musiclab.chromeexperiments.com/Spectrogram)

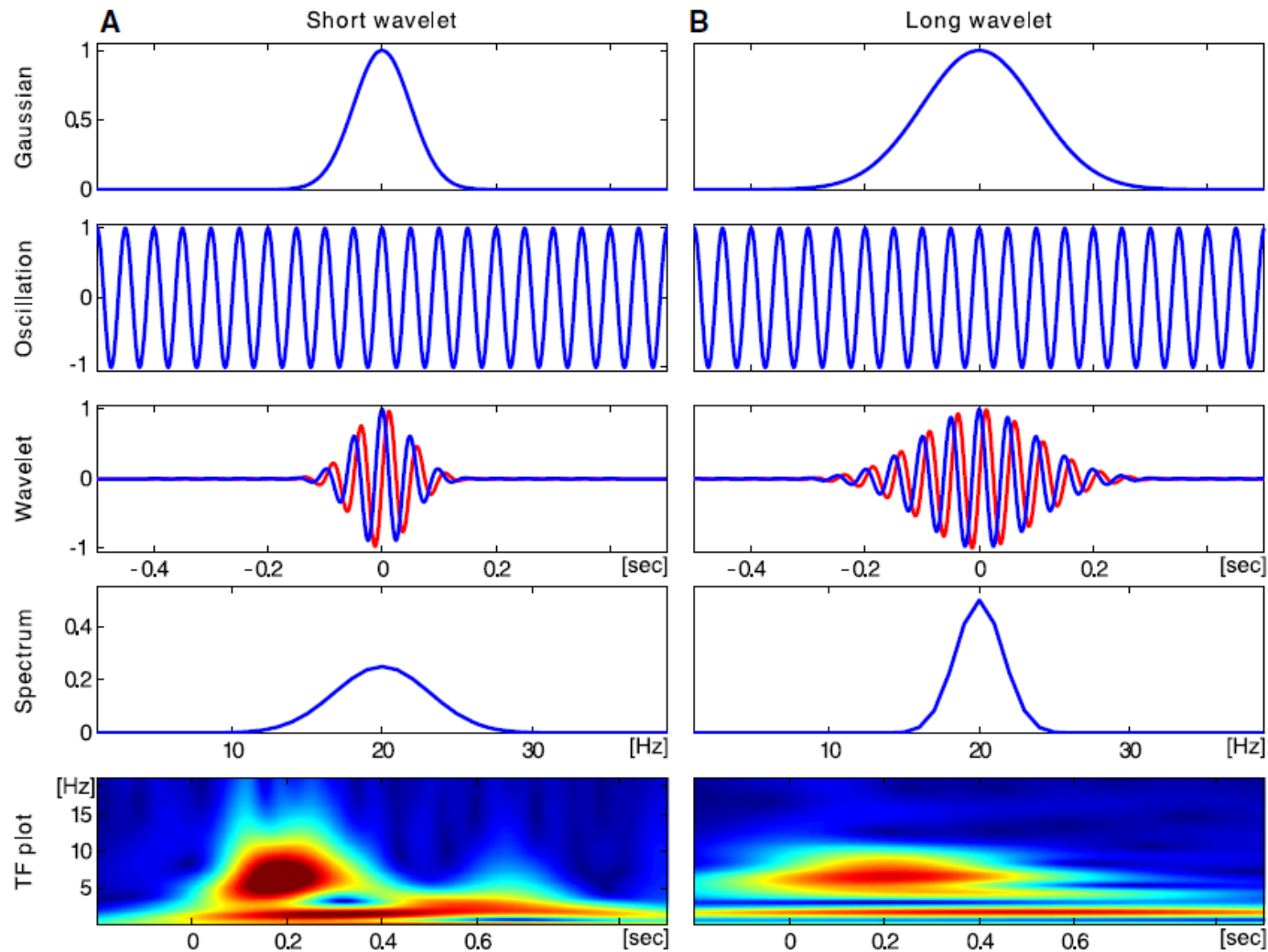
Time-Frequency Approaches

Brain Topogr (2014) 27:438–450



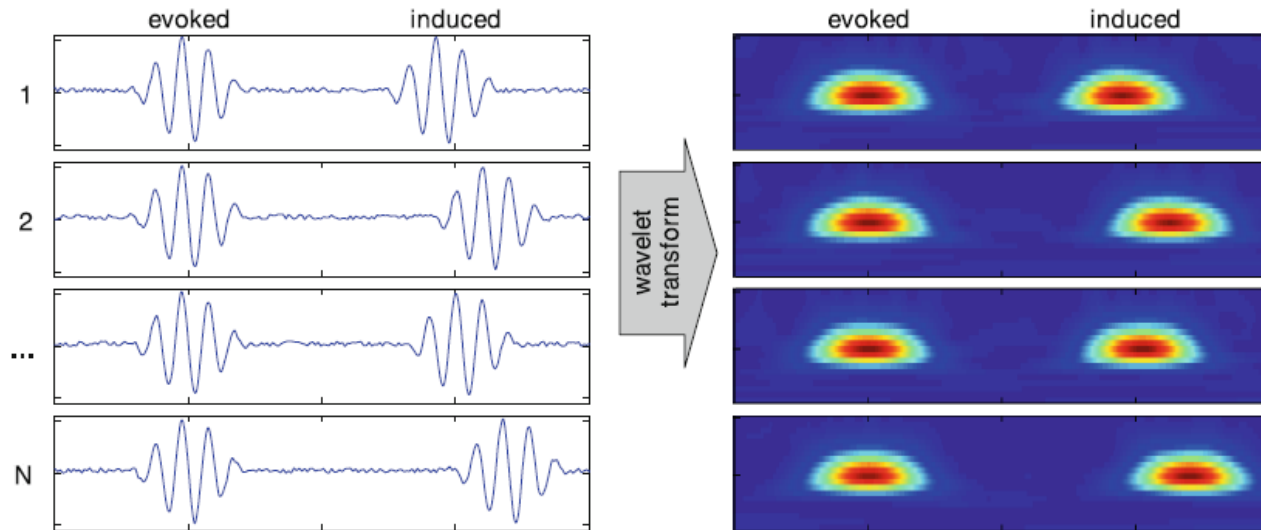
Time-Frequency Approaches

Brain Topogr (2014) 27:438–450



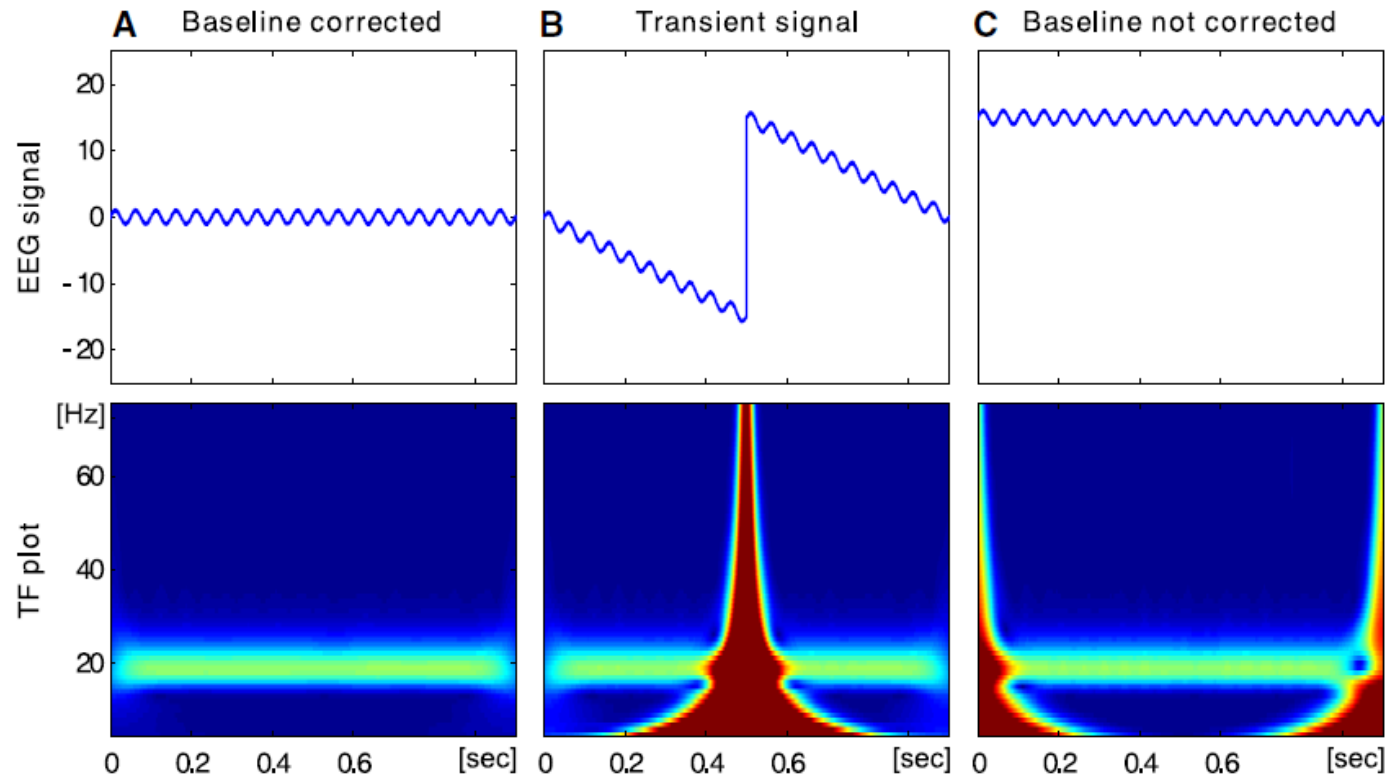
Time-Frequency Approaches

Brain Topogr (2014) 27:438–450



Time-Frequency Approaches

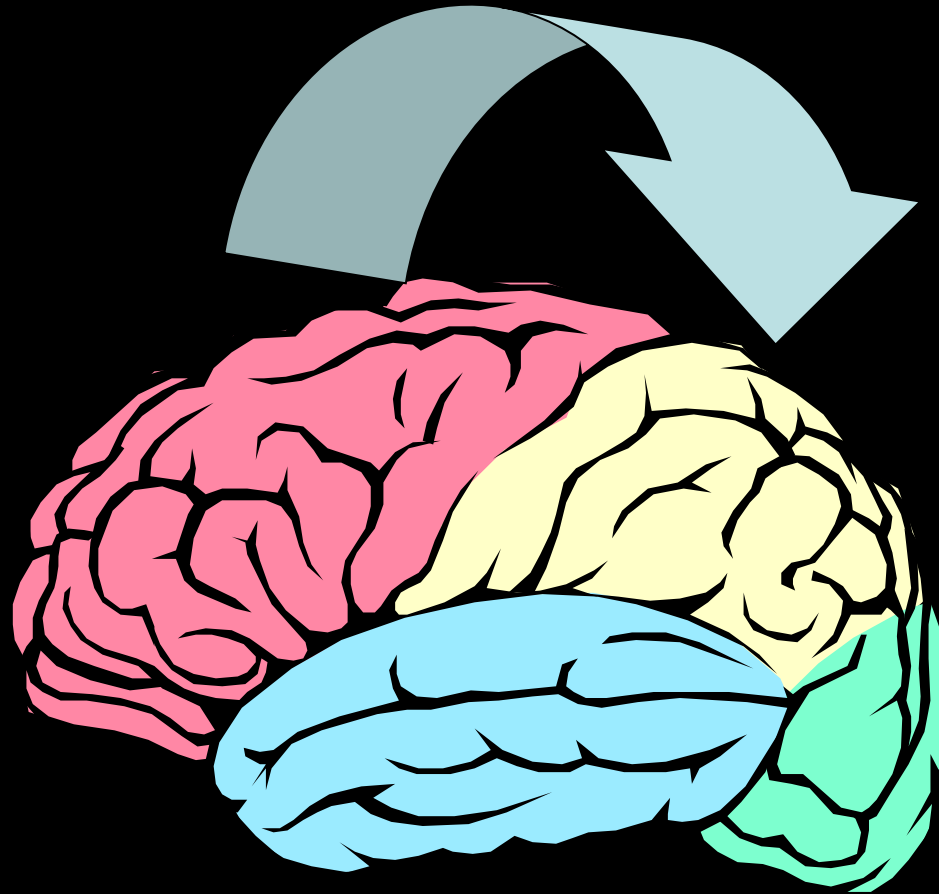
Brain Topogr (2014) 27:438–450



A bit more on phase and such

COURTESY OF MIKE COHEN

2. How do brain regions “talk” to each other?

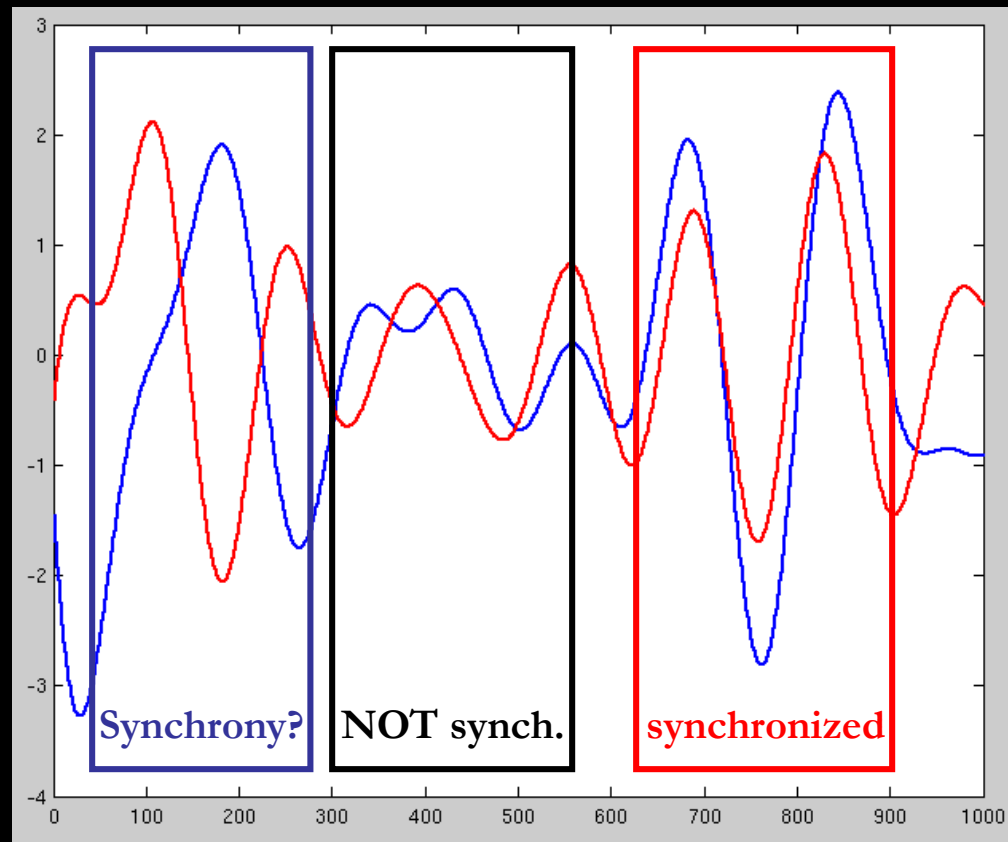


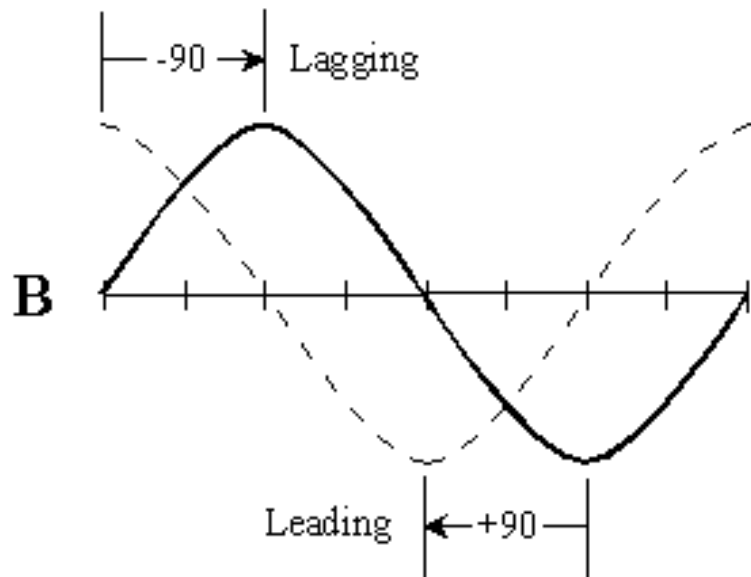
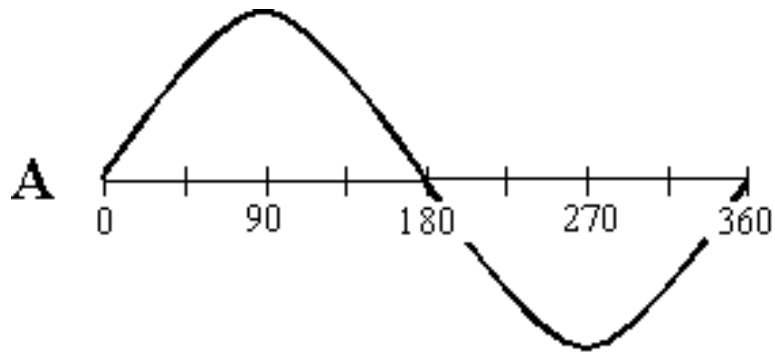
Perhaps through synchronized oscillations!

See empirical work and reviews by:
Rubino, Lisman, Singer, Engels, etc.

2. How do brain regions “talk” to each other?

**Synchronized oscillations is an intuitive concept,
but how to measure it quantitatively?**





➤ The time interval for one degree of phase is inversely proportional to the frequency.

➤ You know.... the frequency of a signal f is expressed in Hz)

➤ The time t (in seconds) corresponding to:

➤ one degree of phase is:

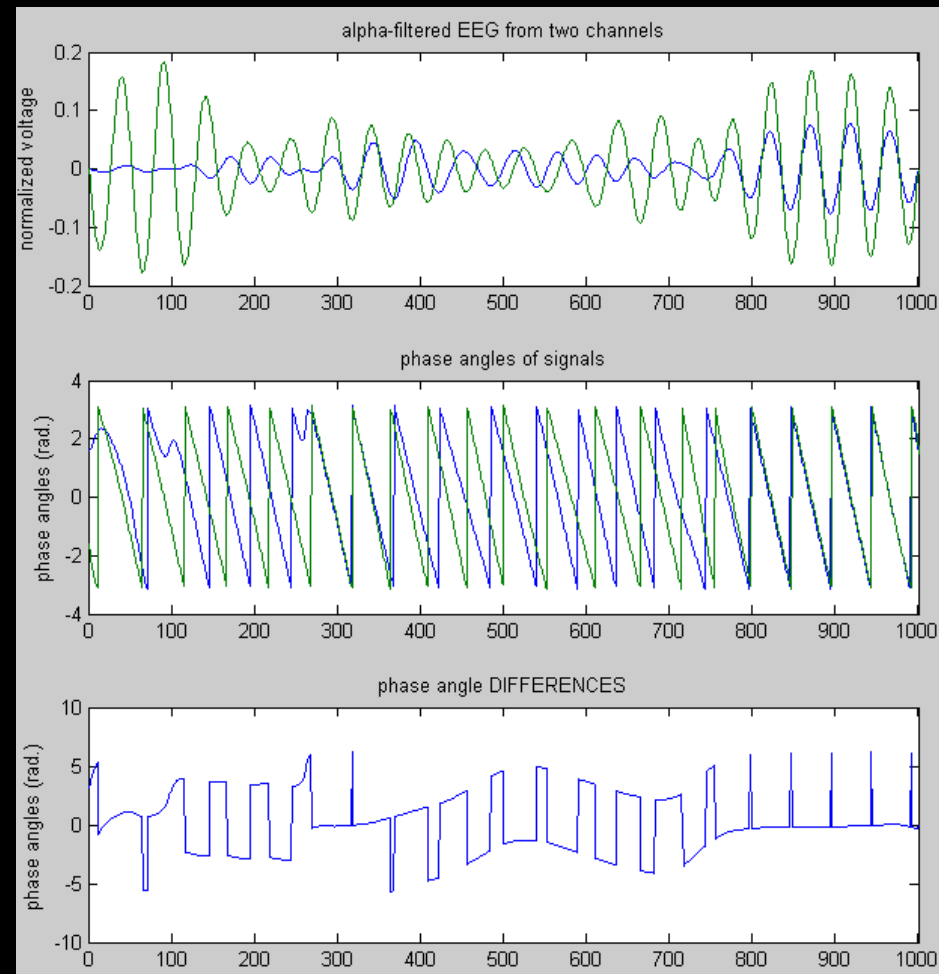
$$t_{\text{deg}} = 1 / (360 f)$$

➤ one radian of phase is approximately:

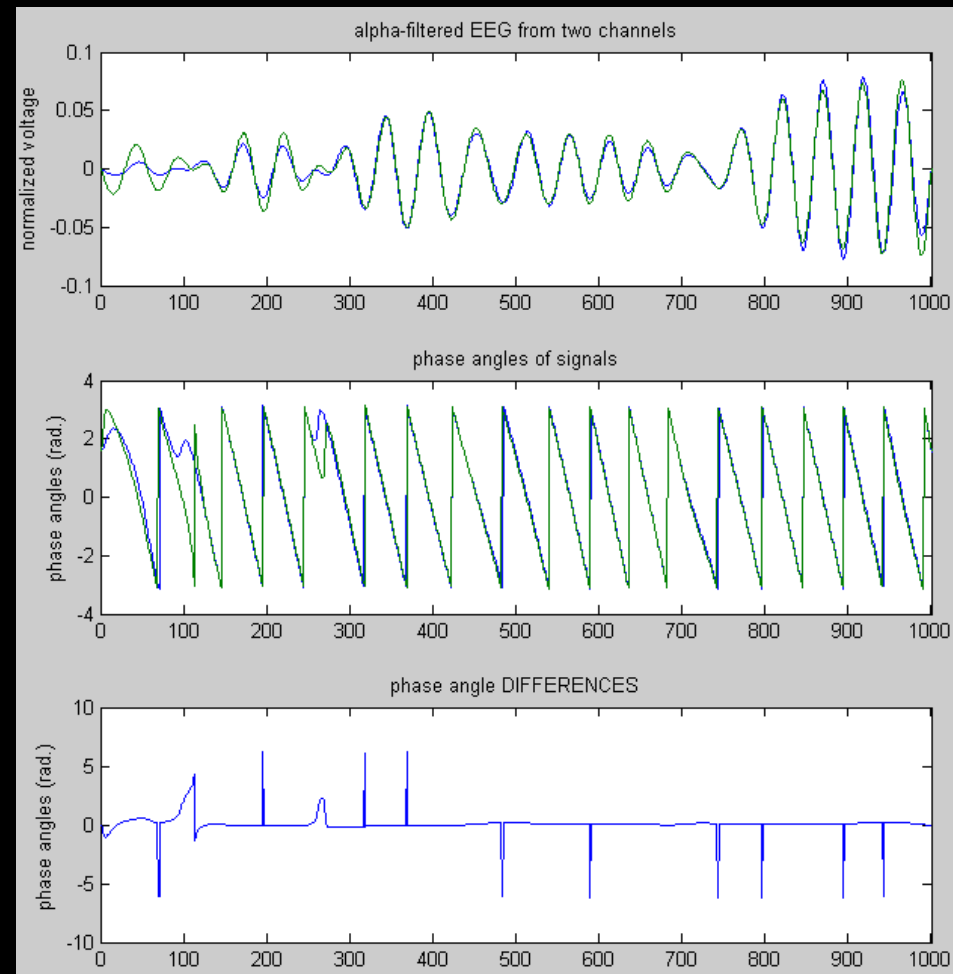
$$t_{\text{rad}} = 1 / (6.28 f)$$

2. Inter-site phase coherence.

Electrodes: Fp1 & C4

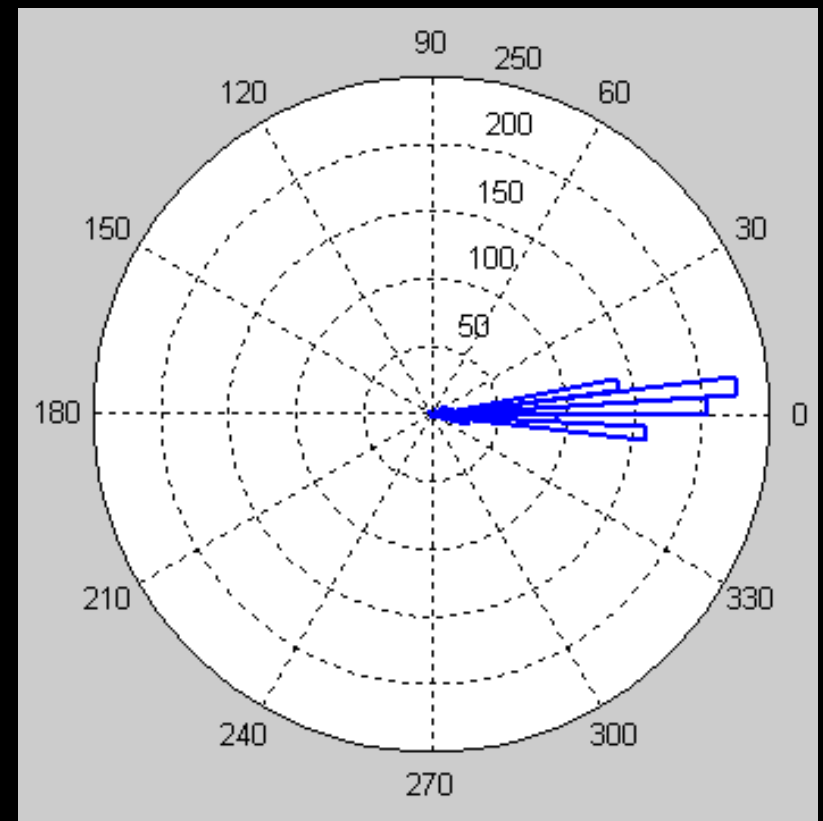
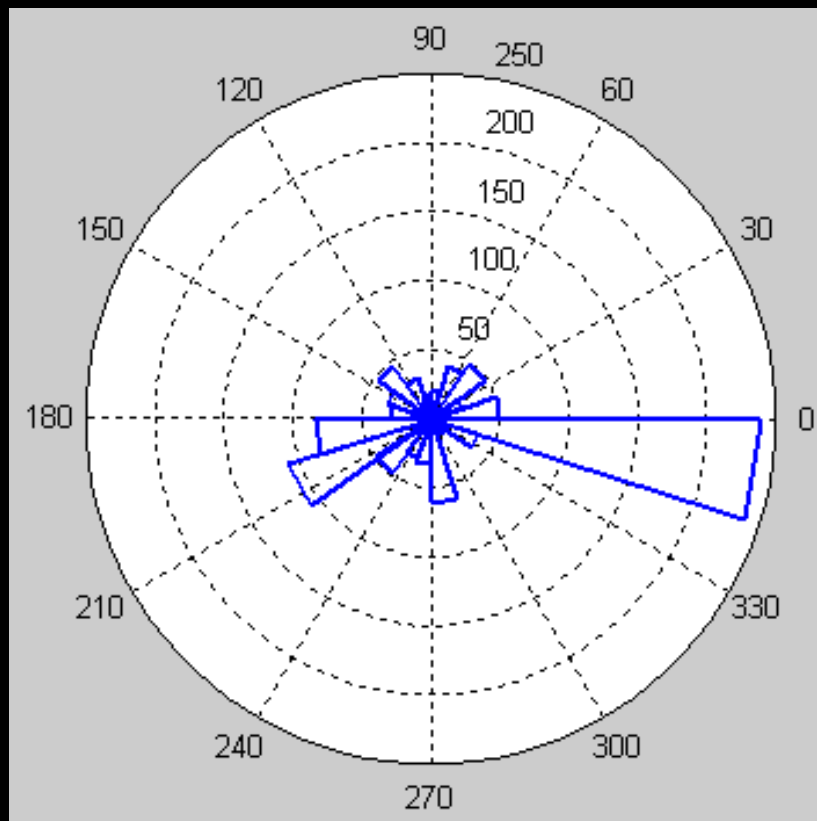


Electrodes: Fp1 & Fp2



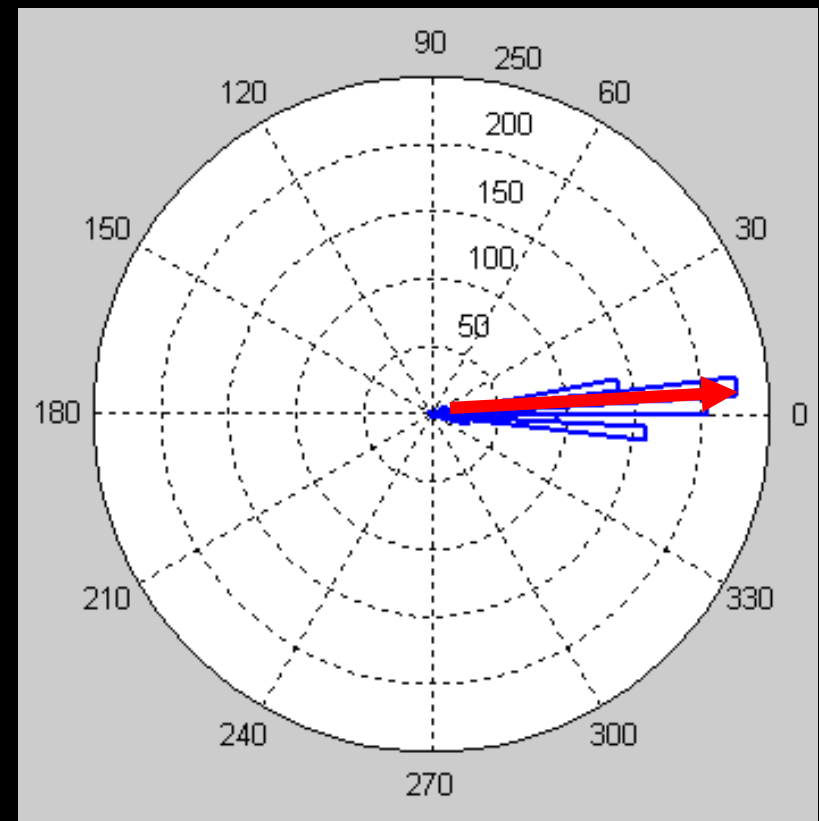
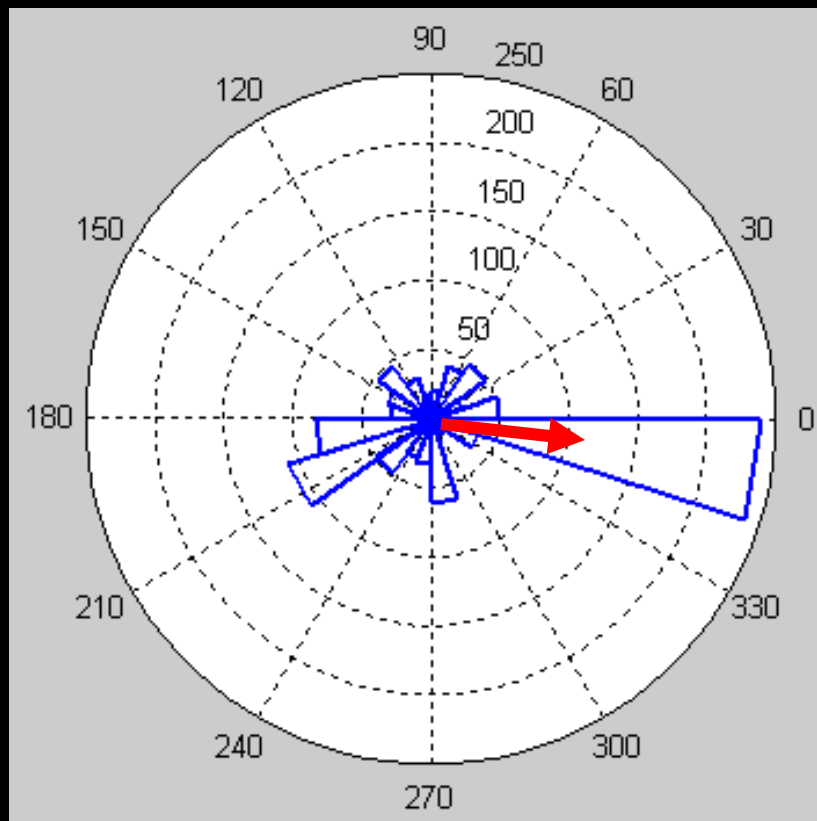
2. Inter-site phase coherence?

“Polar plot” of phase angle differences.



2. Circular variance.

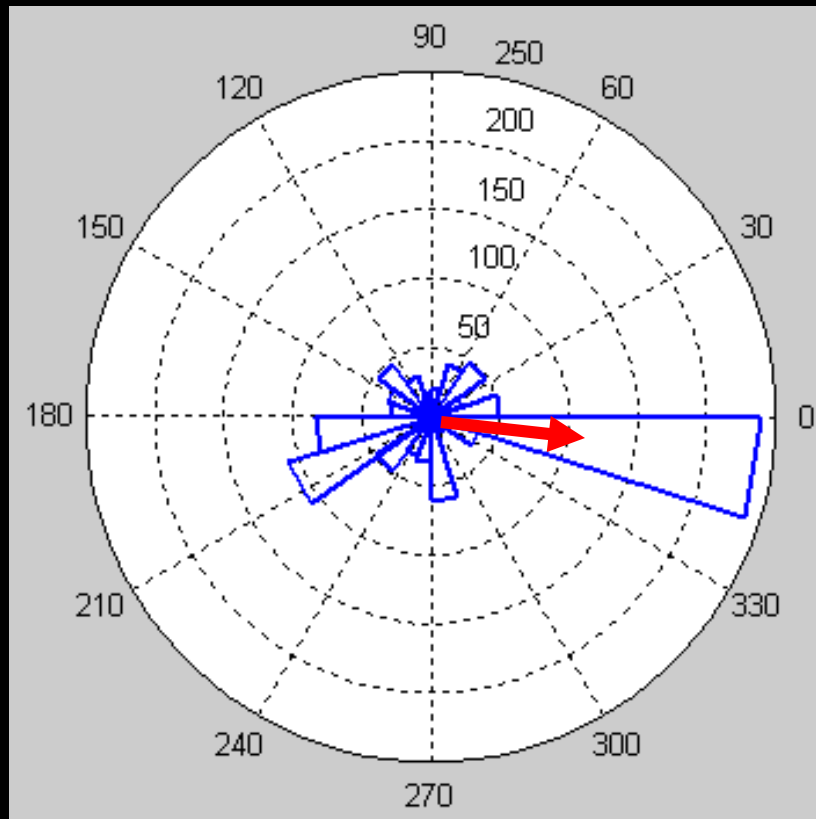
Draw a line through the “average” of vectors.



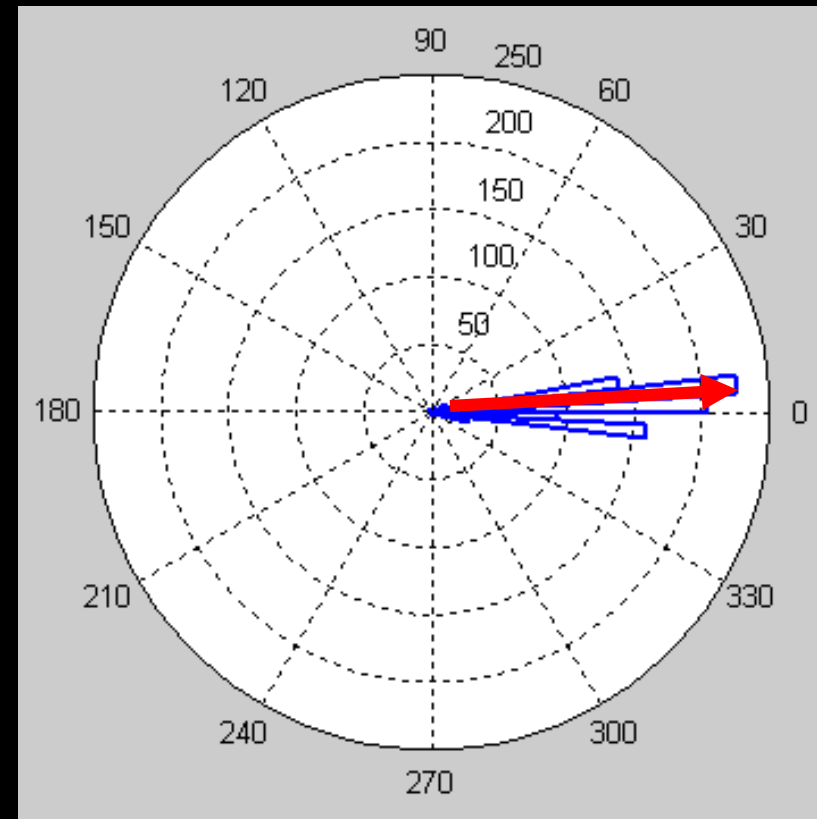
2. Circular variance.

The length (magnitude) of that vector varies from 0 to 1, and is the phase coherence.

Phase coherence: 0.11



Phase coherence: 0.94



2. Circular variance.

The equation for phase coherence is simple:

```
> abs(mean(exp(i*angle_differences)) );
```

**Magnitude
of vector**



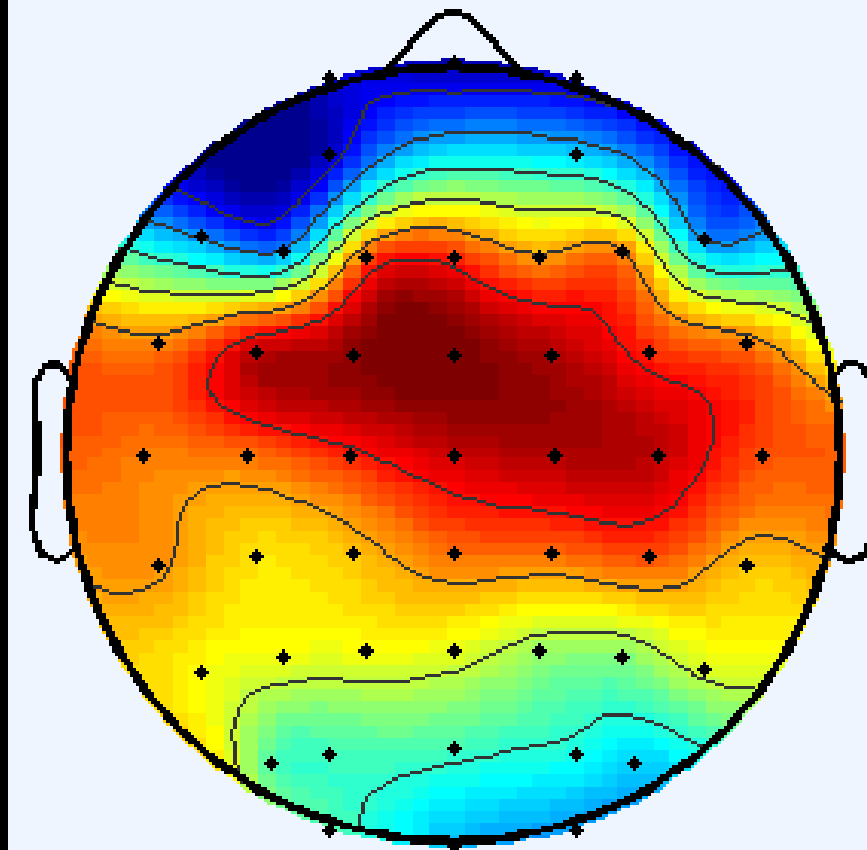
**Average
across
values**

**Transform to
complex plane**

**Phase angle
differences
between
channels**

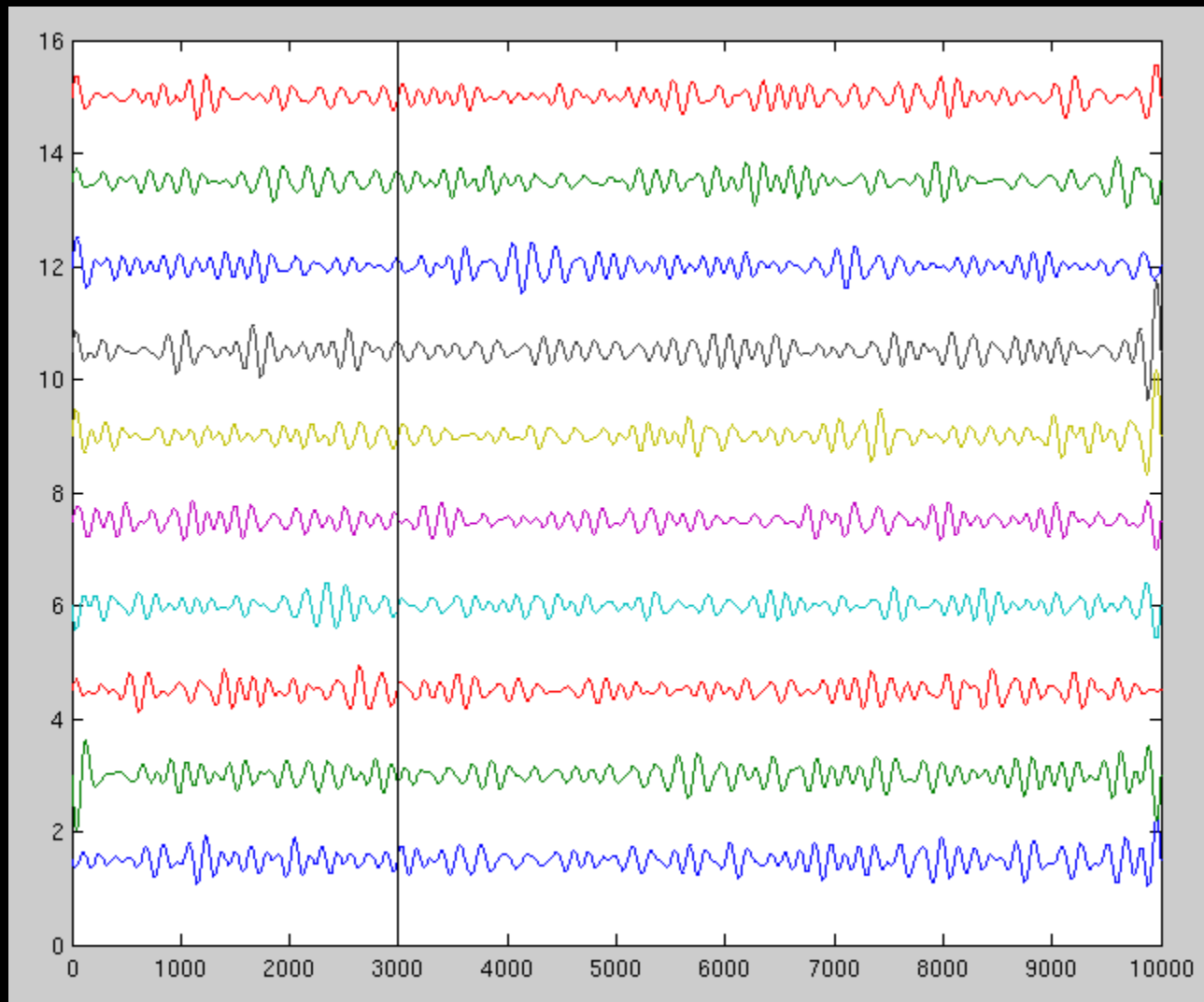
2. Inter-site phase synchrony with one “seed” site.

Phase coherence with channel: FCZ

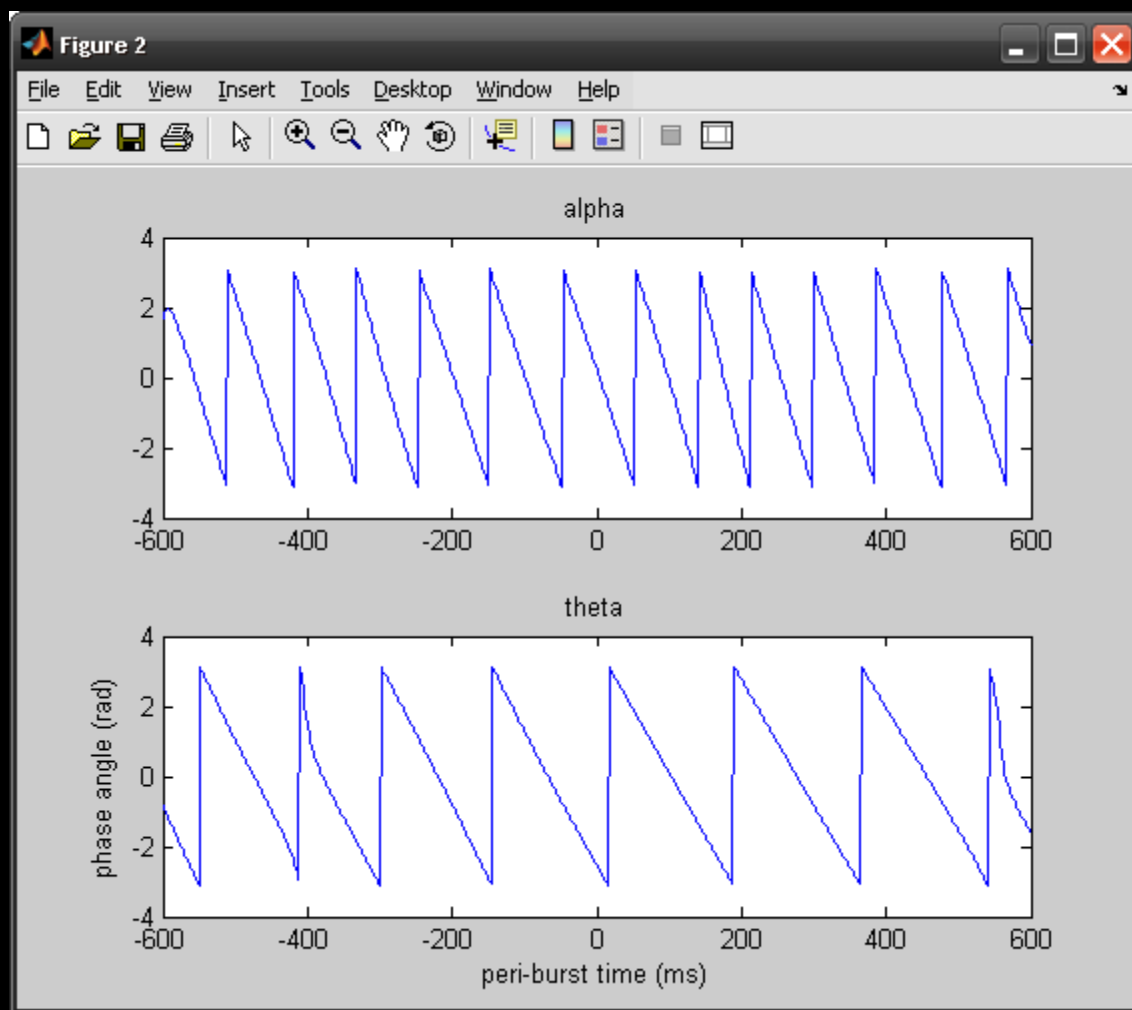


2. Inter-trial phase synchrony within one electrode.

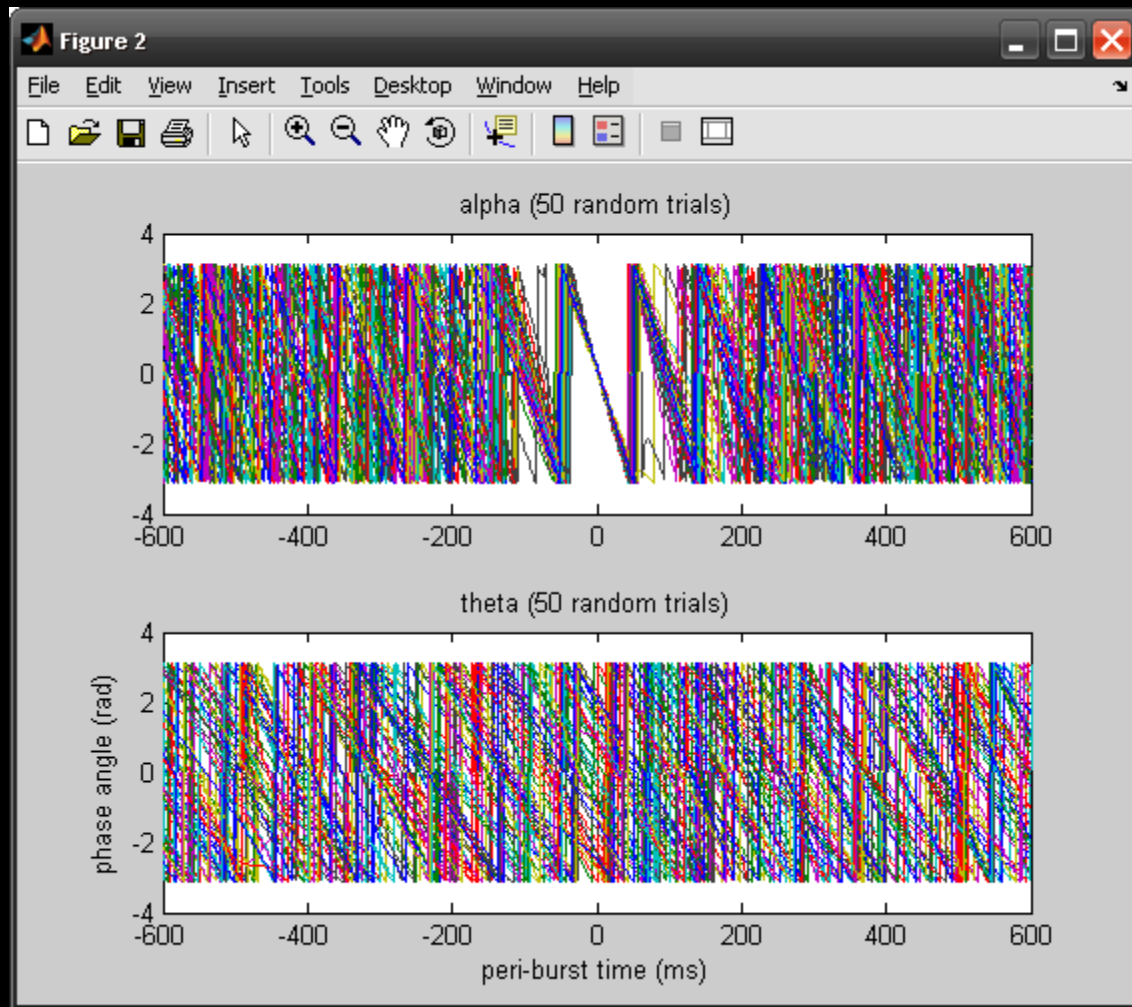
Many trials from the same electrode:



2. Inter-trial phase coherence



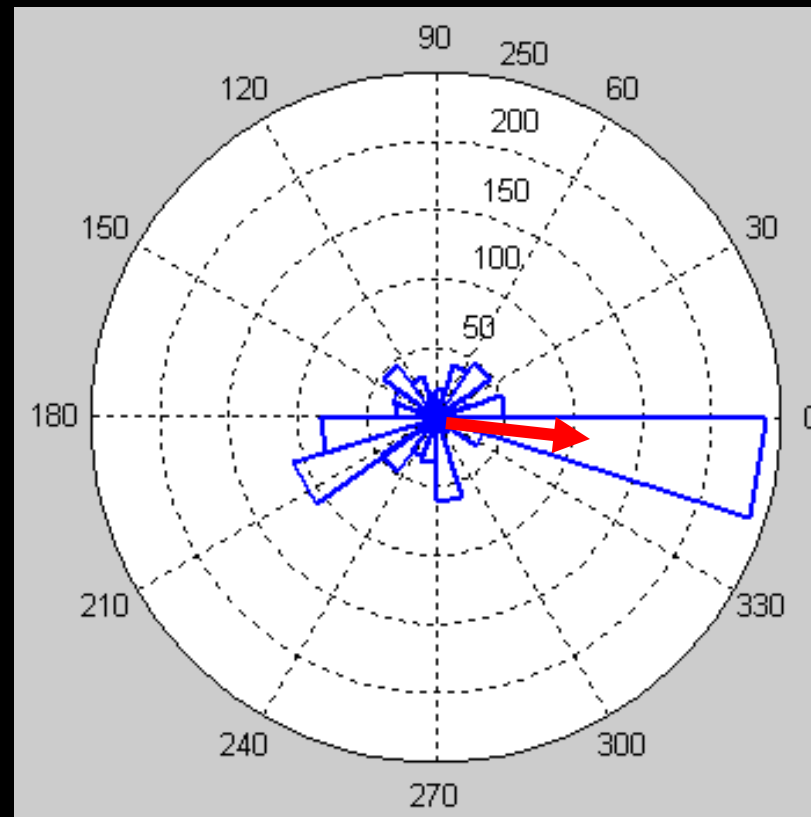
2. Inter-trial phase coherence



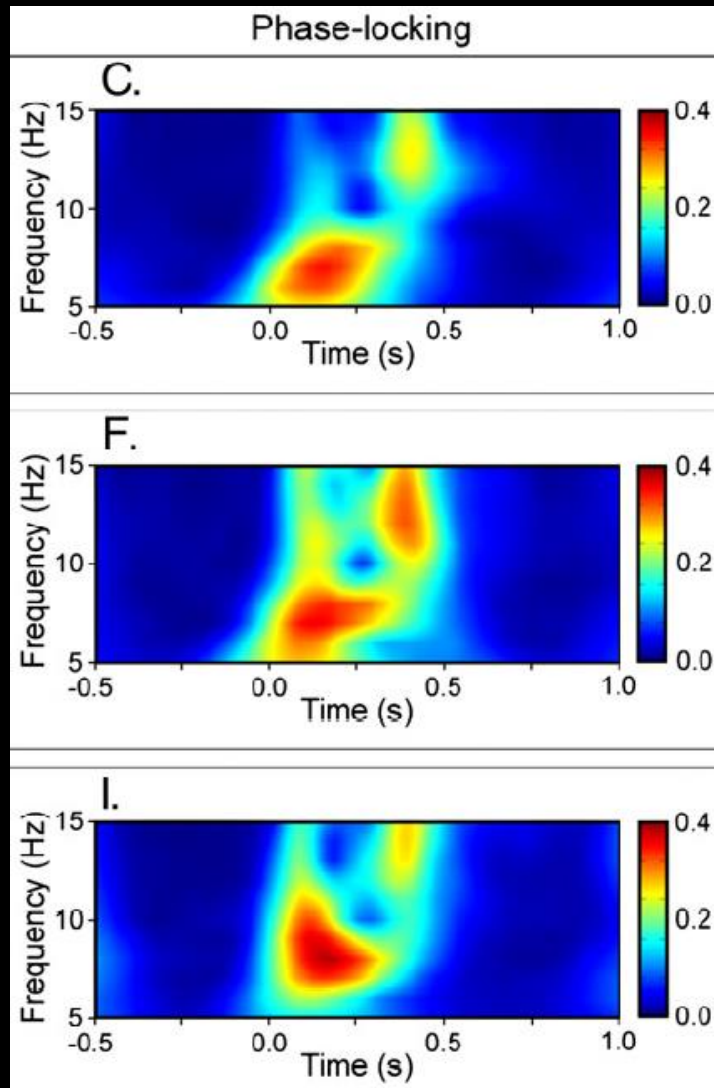
2. Inter-trial phase coherence

Calculate phase coherence across trials at each time point

Phase coherence, 154 ms: 0.11



2. Inter-trial phase coherence



3 different electrodes

Thanks Mike!

NOW BACK TO JOHN'S SLIDES

Power increase in the absence of any phase locking

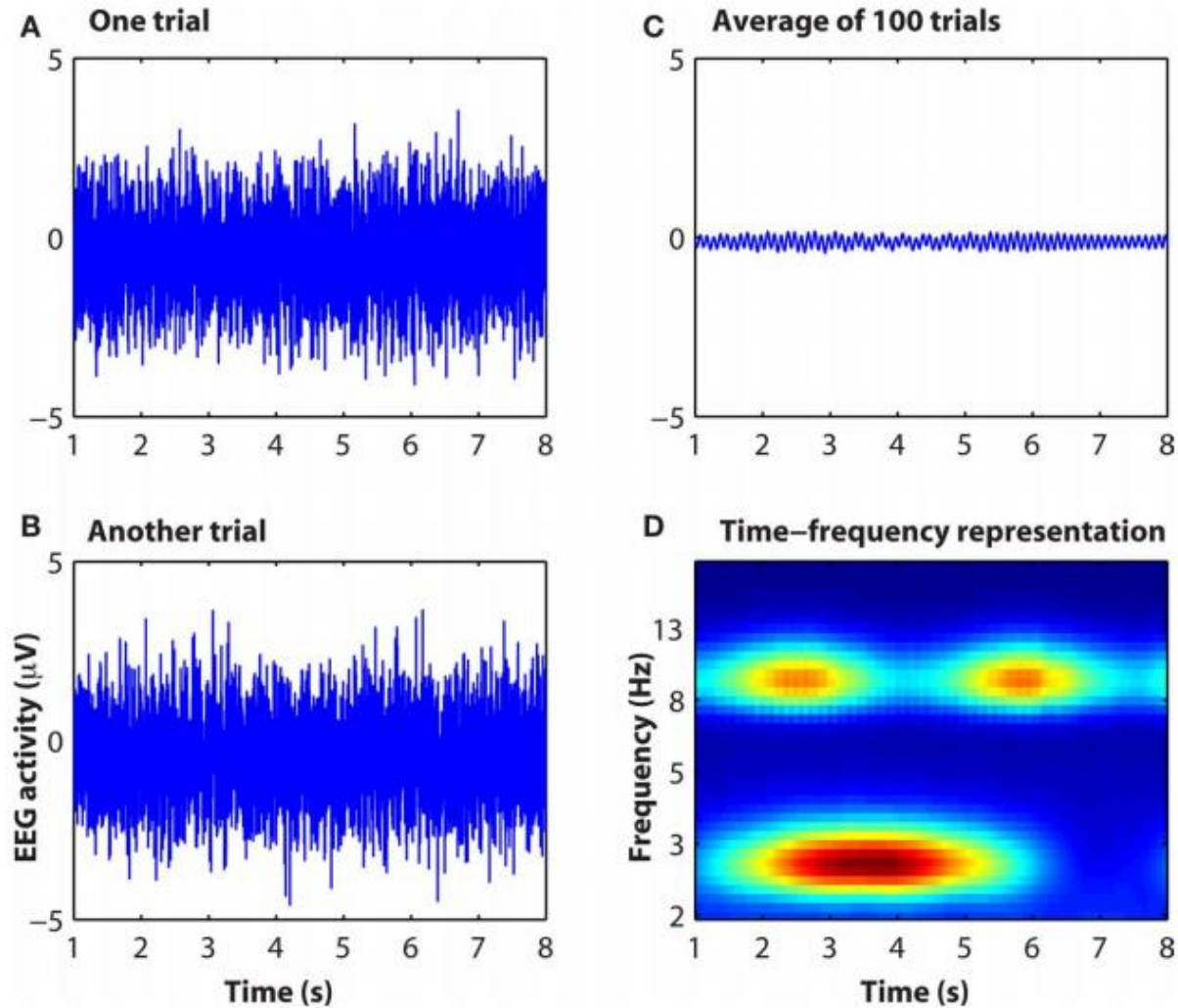


FIGURE 3 | Simulated data showing how information contained in raw EEG data [(A,B): single “trials”] is not apparent in the event-related potential (C) but is readily observable in the time–frequency representation (D). Matlab code to run this simulation is available from the author.

The Importance of Phase!

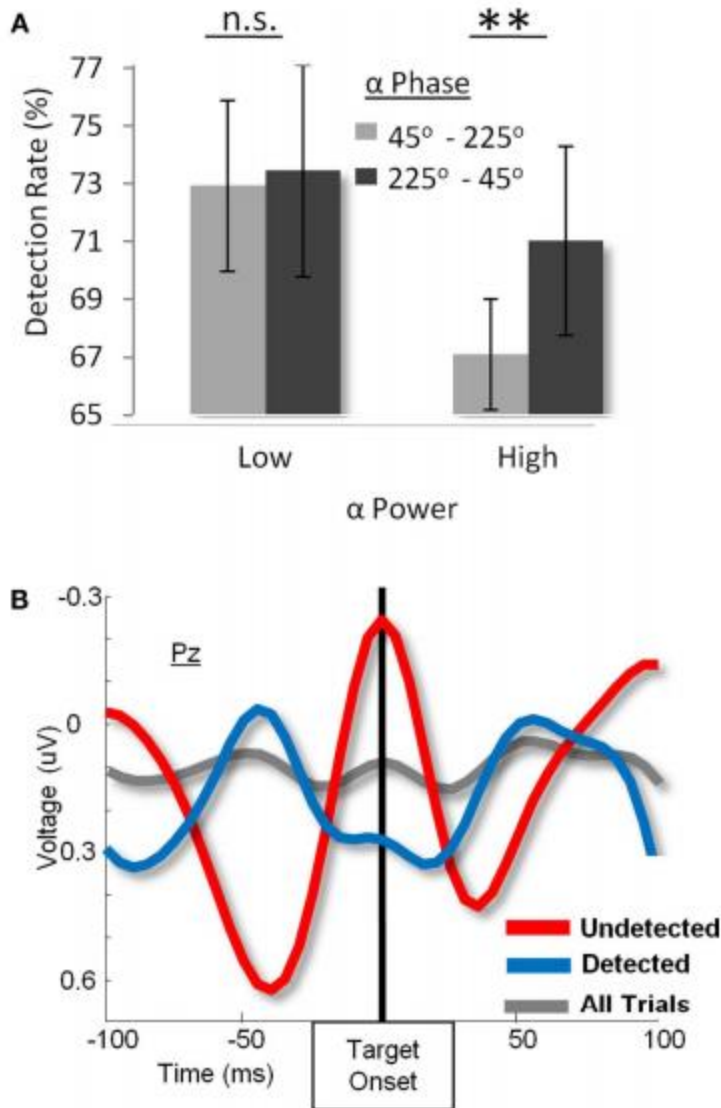
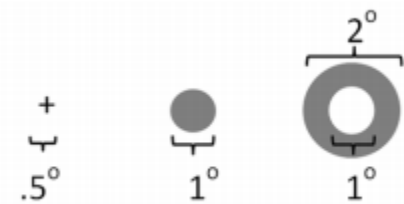
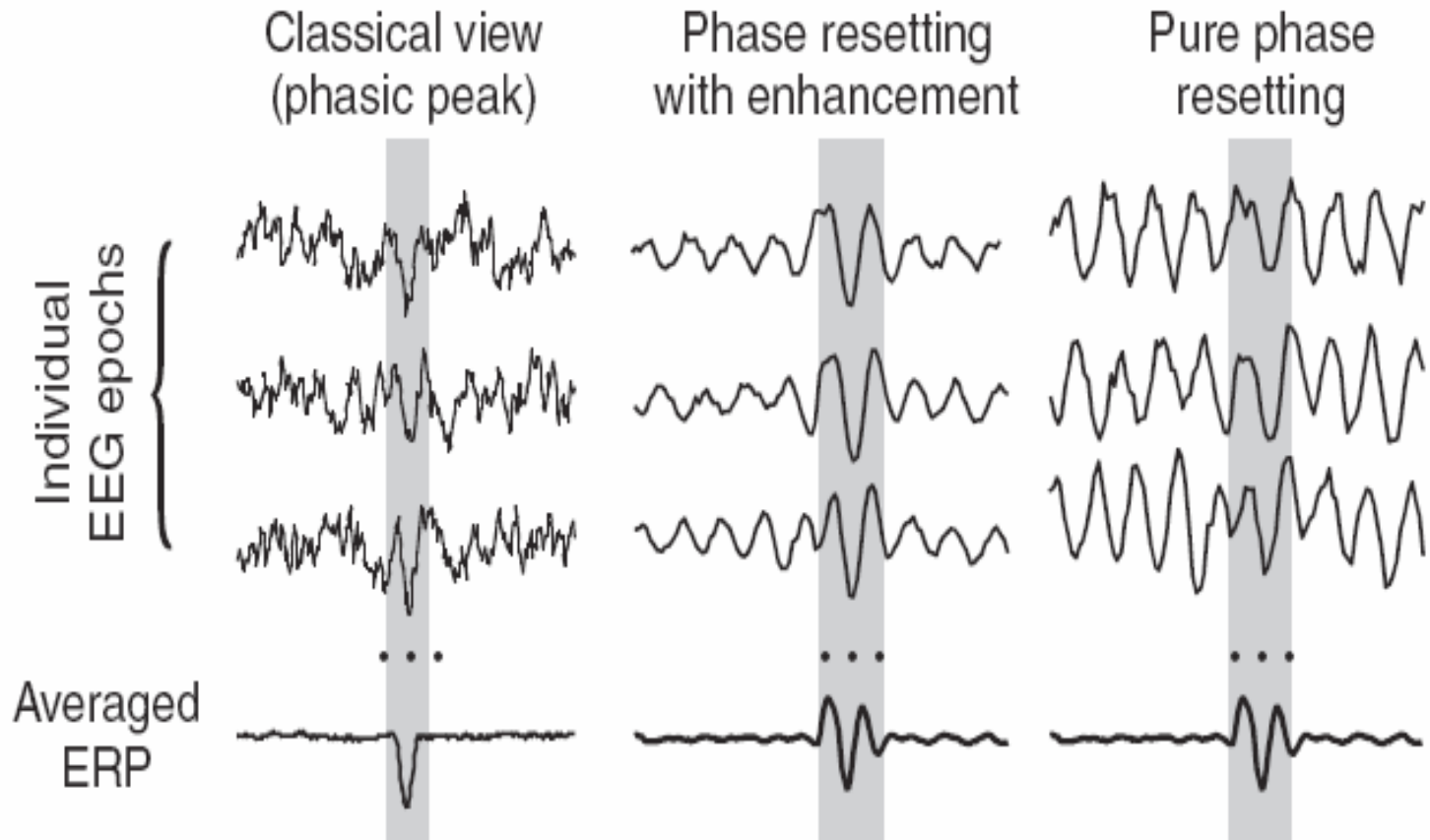


FIGURE 3 | (A) Detection rate as a function of alpha power and phase before stimulus onset. When alpha power is low (left bar graph), there is no difference in masked-target detection as a function of pre-target alpha-phase. When alpha power is high (right bar graph), however, not only is detection lower overall, but it differs between opposite alpha-phases. **(B)** Grand-average ERP at the Pz electrode for detected (blue), undetected (red), and all (gray) targets. Results show the presence of counterphase alpha oscillations between detected and undetected targets, whereas the overall average is flat, indicating that subjects did not phase lock to the stimulus before its onset. **(C)** Polar plot of a bootstrap-derived distribution of the average phase (angle) and amplitude (distance from origin) of pre-target 10-Hz oscillations for detected (red) and undetected (blue) targets. Each dot is the grand-average phase over the 12 subjects for one of 10,000 equally sized random samples from the two conditions. The arrows represent the centroids of the distribution of mean phases. (Figure adapted from Mathewson et al., 2009, reprinted with permission).

Time-Frequency Approaches to Error Monitoring

Classic ERPs Vs Phase Resetting



Time-Frequency Representations

648

L.T. Trujillo, J.J.B. Allen / *Clinical Neurophysiology* 118 (2007) 645–668

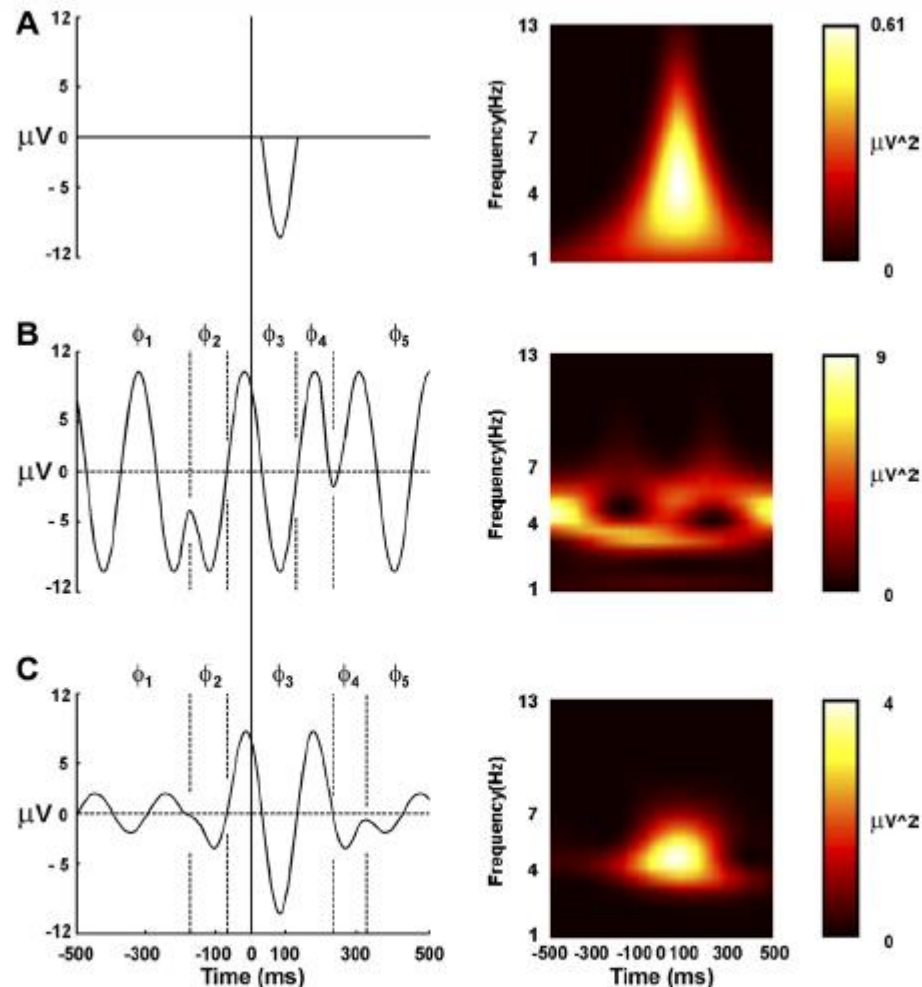
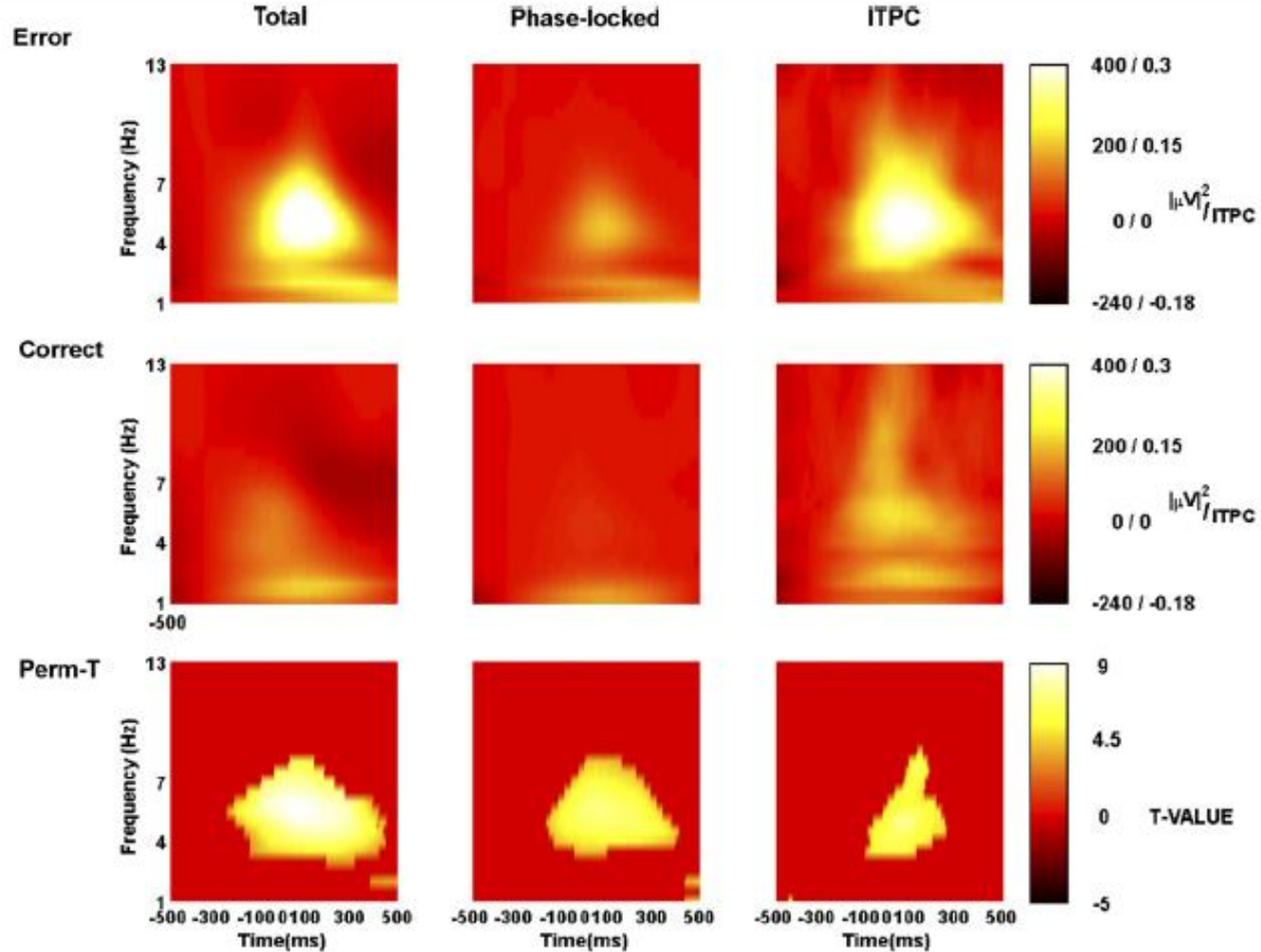
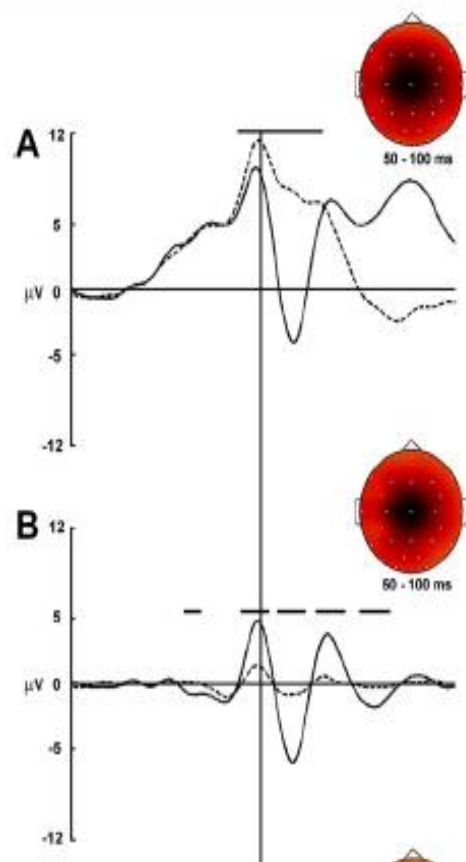
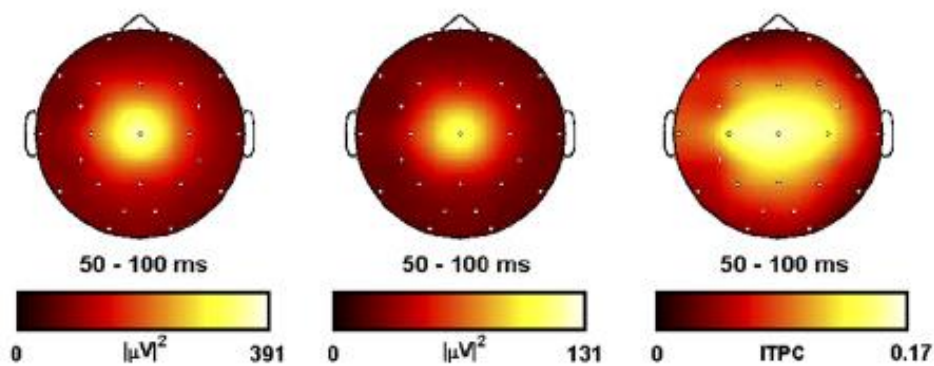


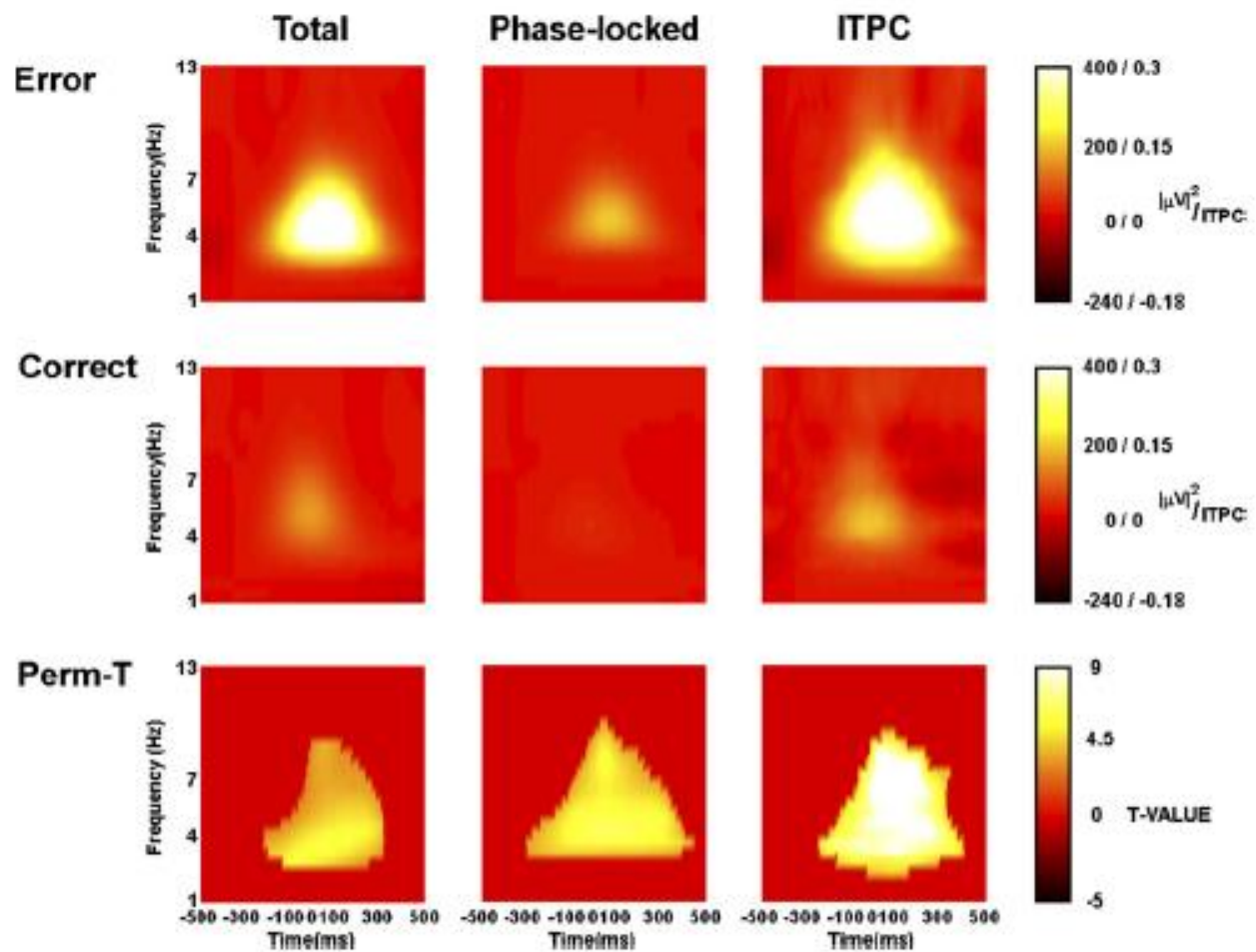
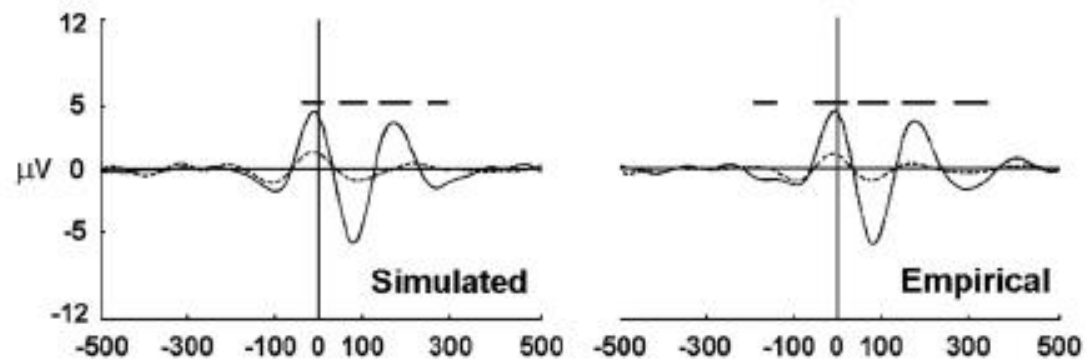
Fig. 1. Left column: Basic oscillatory waveforms used to simulate ERN responses according to the (A) *classic*, (B) *pure phase-resetting*, and (C) *phase-resetting with enhancement* hypotheses of ERN generation. Right column: Corresponding non-baseline-corrected wavelet-based time-frequency representations of these waveforms. The procedures used to create these waveforms and time-frequency representations are described in Sections 2.6 and 2.7.



Topo

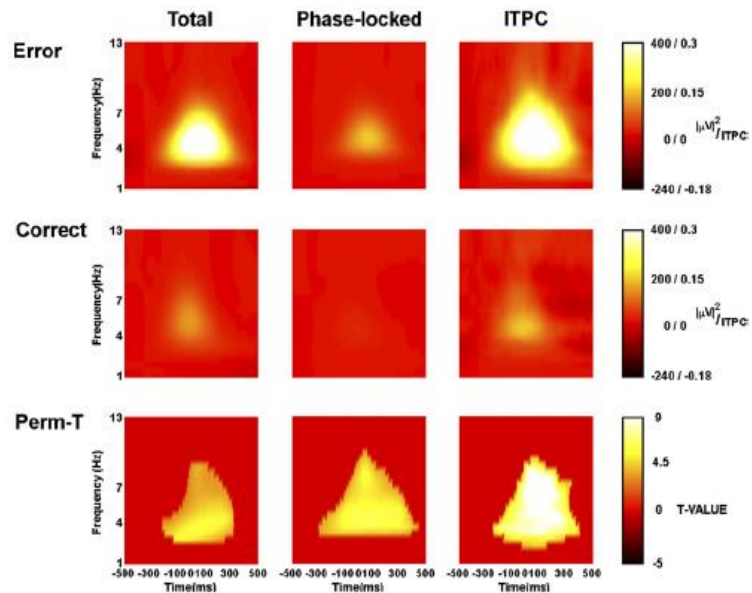
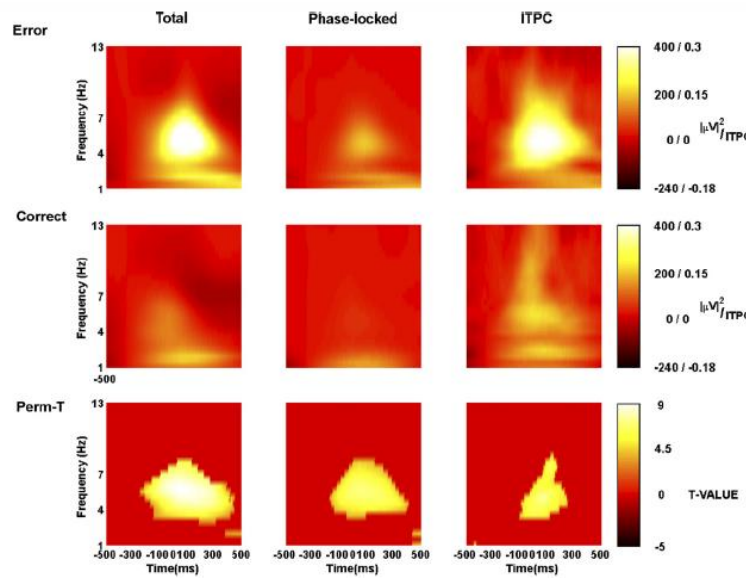


Simulated Phase-resetting with Enhancement

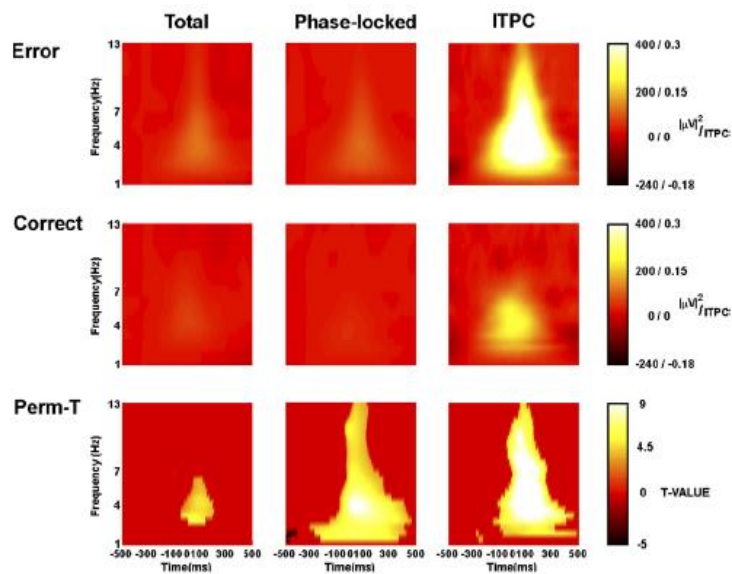


Empirical

Simulated Phase + Amp Enhance



Simulated Classic



Dealing with Ocular Artifacts

Ocular Artifacts

- The problem
 - Eye movements and blinks create a potential that is propagated in volume conducted fashion
 - Manifests in recorded EEG
- Why?
 - Eye not spherical; more rounded in back
 - Potential is therefore positive in front with respect to rear of eye
 - Movements = Moving dipole
 - Blinks = sliding variable resistor

Ocular Arifacts

- Eye-blinks are *systematic* noise with respect to the ERP signal
 - Occur at predictable latencies (Stim-Resp-Blink)
 - Are meaningful variables in and of themselves:
 - John Stern: Information processing and blink latency
 - Peter Lang: Blink Amplitude and affectively modulated startle response

Ocular Artifacts

- Signal averaging will not remove this "noise" (noise wrt signal of interest)
- Average waveform $a(t)$ is mixture of time-locked signal $s(t)$ and randomly distributed error (noise)

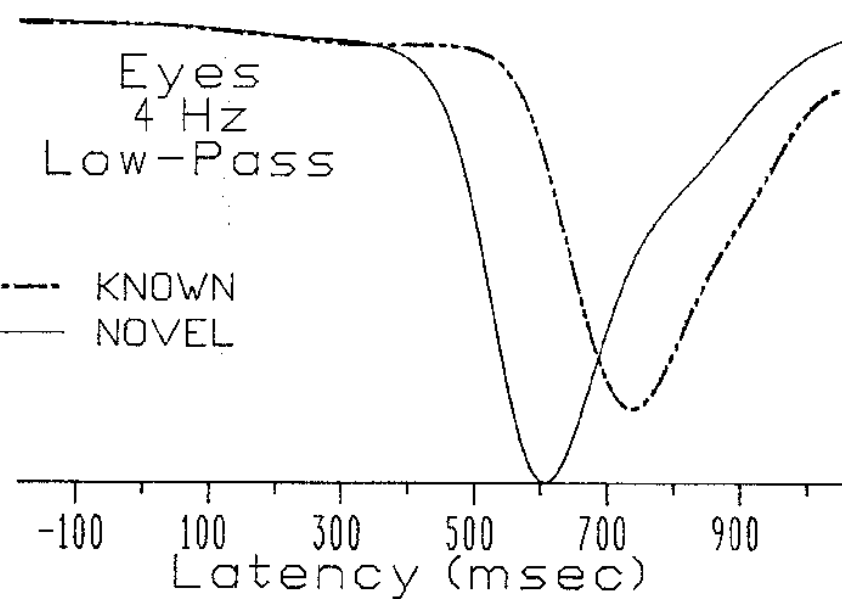
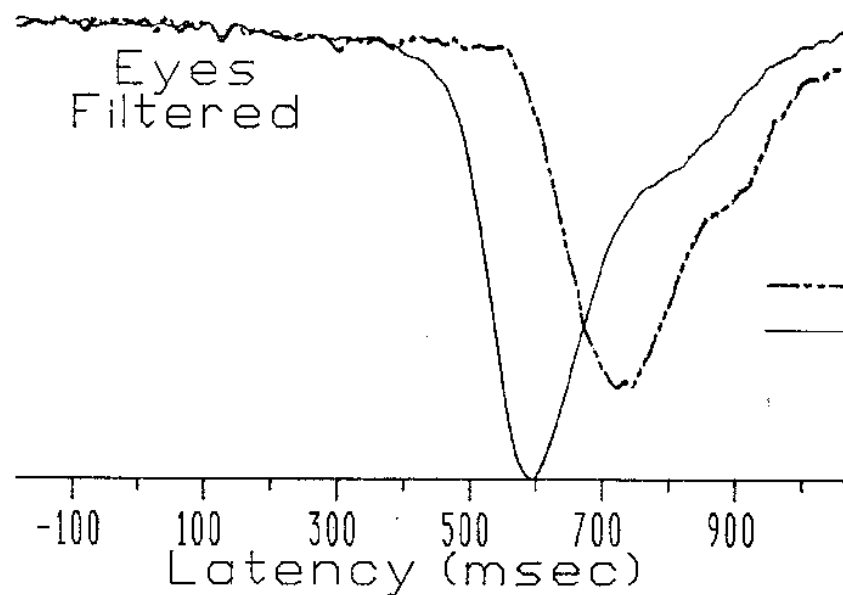
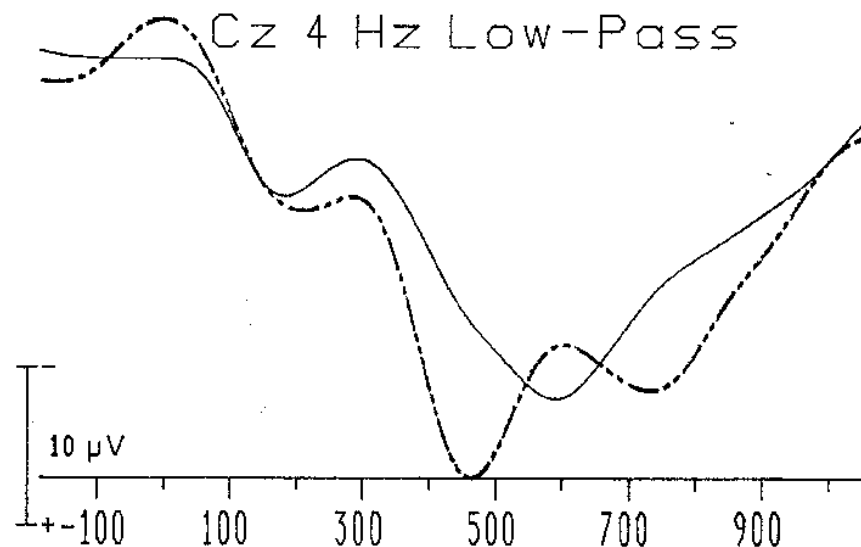
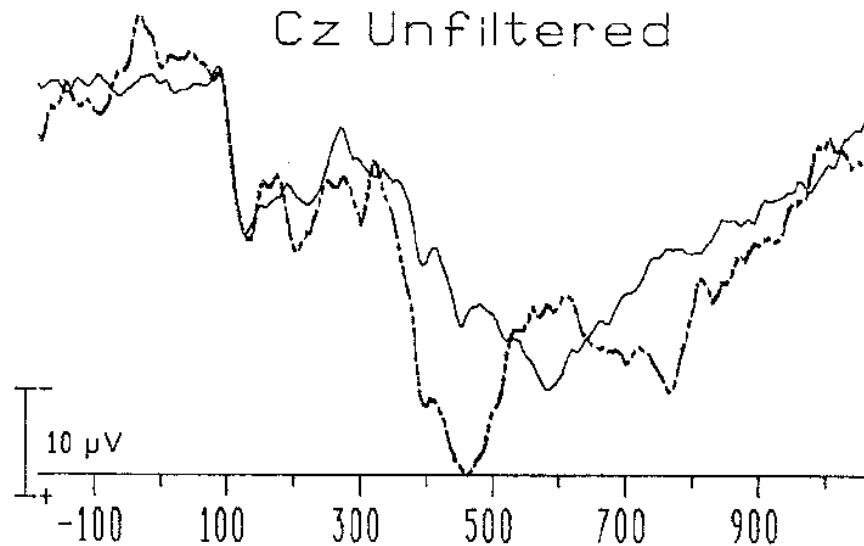
$$a(t) = s(t) + \frac{\sum_{n=1}^n e(t)}{n}$$

- If non-ERP signals are random with respect to stimulus onset, then the latter term will approach zero with sufficient trials (n)
- If not, the latter term will not sum to zero, but will include time-locked noise
- Noise will therefore average IN, not average OUT

Ocular Artifacts

- Eye-blinks tend to occur at the cessation of processing.
 - Recall that the P300 is also a good index of cessation of processing.
- As a result, eye-blink artifact tends to appear as a late P300ish component

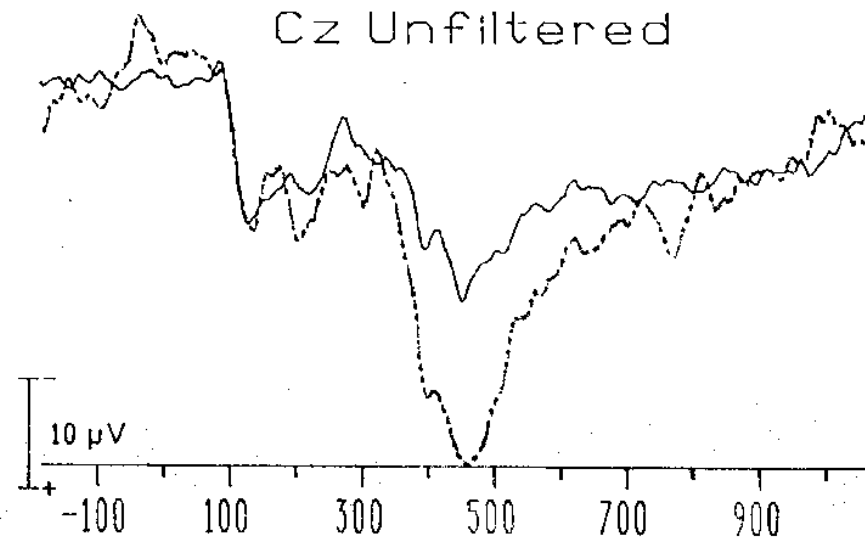
Odd-Ball ERP's SANS Blink Correction



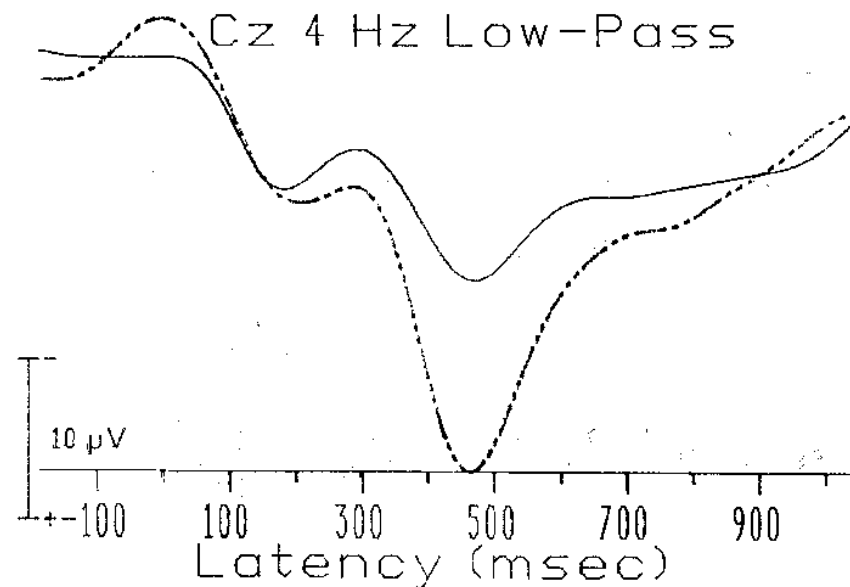
--- KNOWN
— NOVEL

Odd-Ball ERP's WITH Blink Correction

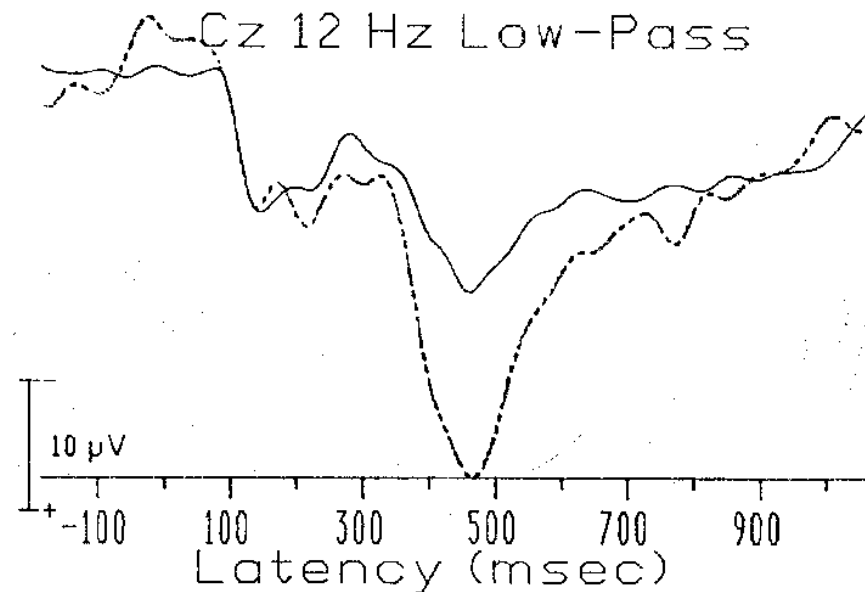
Cz Unfiltered



Cz 4 Hz Low-Pass



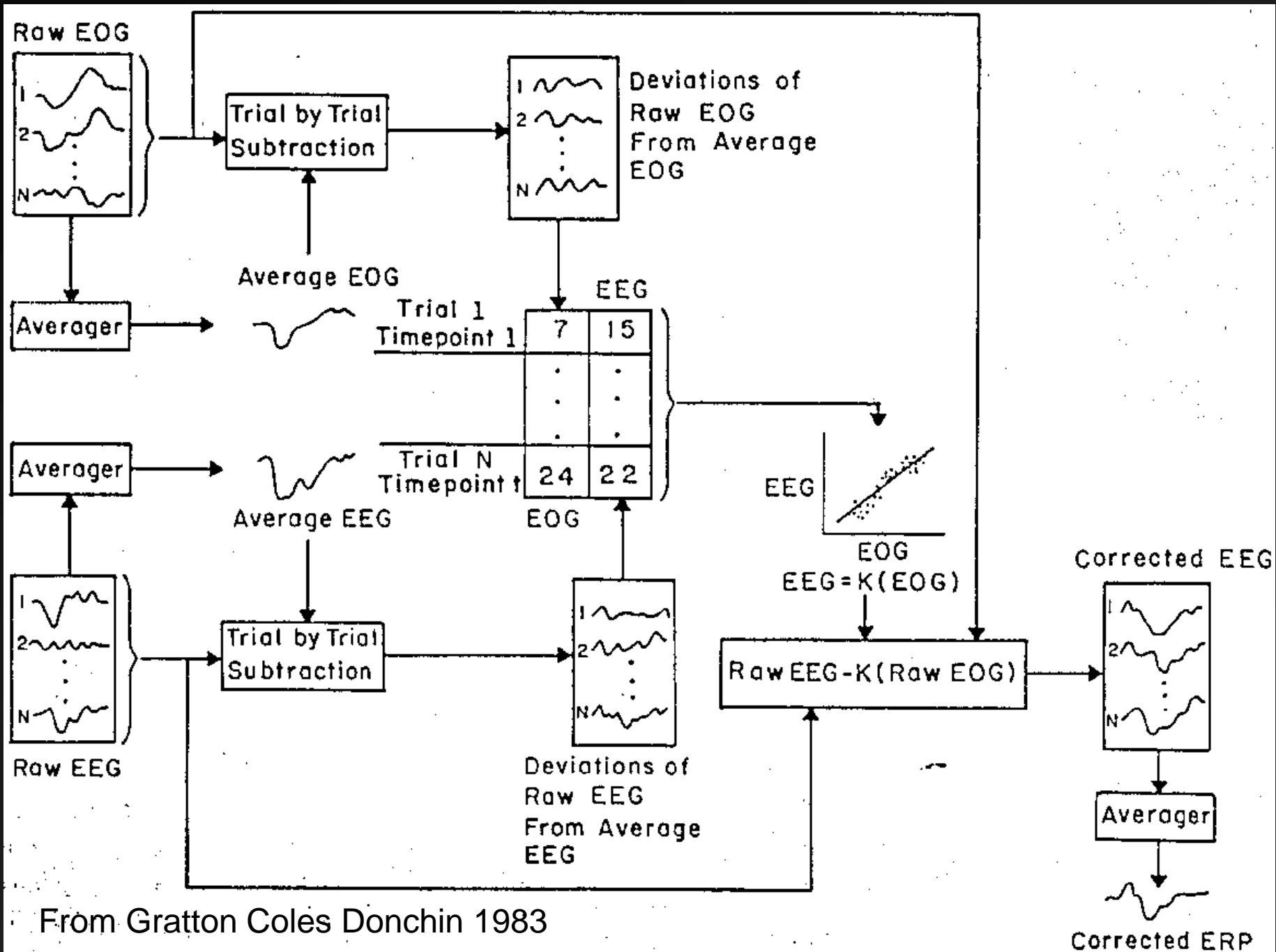
Cz 12 Hz Low-Pass



--- KNOWN
— NOVEL

What to Do?!

- Reject trials during which eye-blink occurred.
 - Problems:
 - Trials which elicit blinks may not be equivalent to those which do not.
 - Large data loss, may be unable to get usable average
 - Telling subjects not to blink creates dual task
- Eye-blink correction (Gratton, Coles, & Donchin, 1983)
 - Assumes that the effect of an eye-movement or blink on the recorded EEG can be inferred from activity recorded near the source of the artifact (top and bottom of eye, e.g.)
- Model ocular potentials as a source, and remove from scalp sites (more later)



The Details

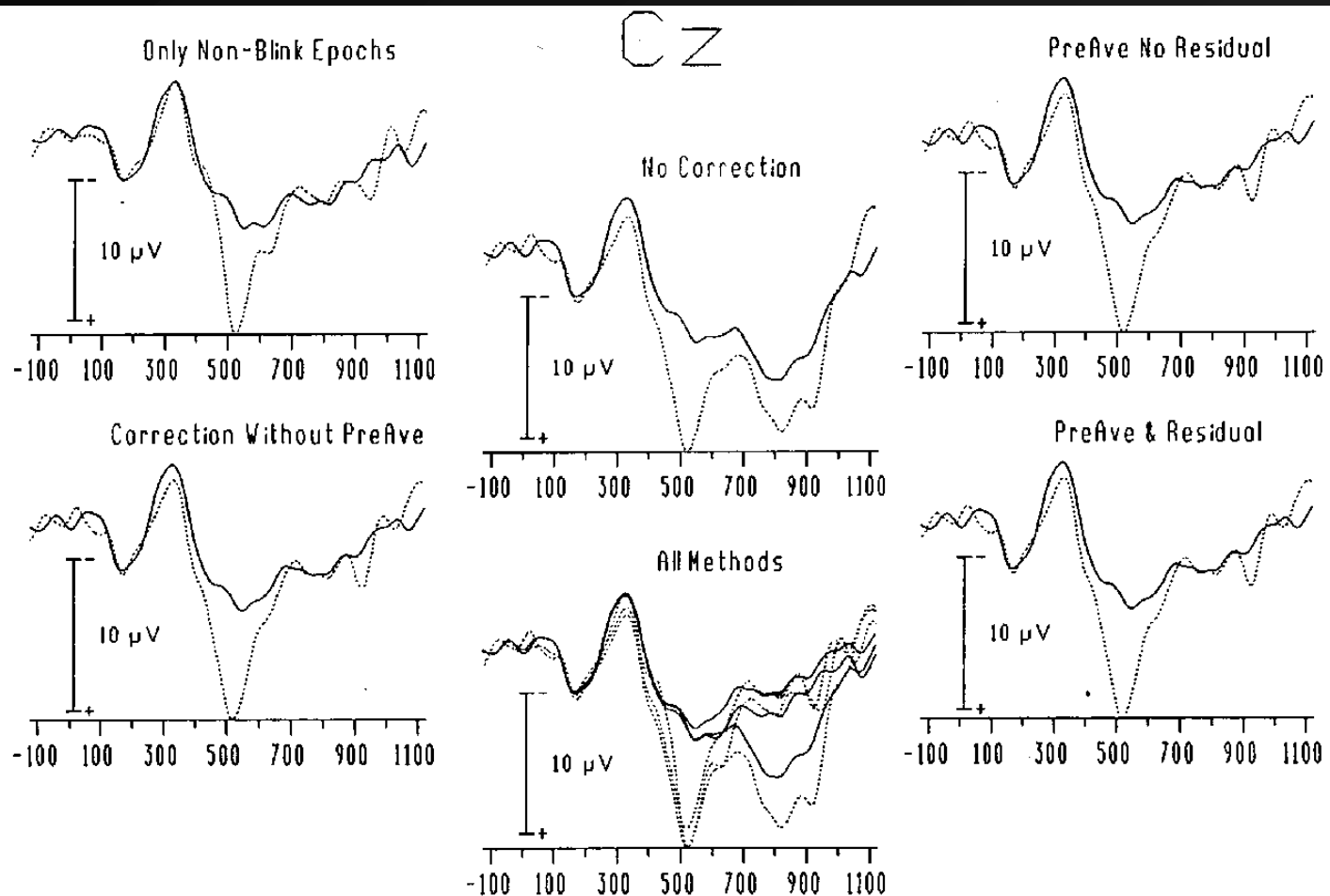
- Must determine extent to which EOG signal propagates to various scalp loci
 - Propagation factors computed only after any event-related activity is removed from both EOG & EEG channels
 - Event related activity in both channels may spuriously inflate estimate of propagation
 - Based upon correlation and relative amplitudes of EEG & EOG, a scaling factor is computed. The scaling factor is then applied on a trial by trial basis as follows:

$$\text{Corrected EEG} = \text{Raw EEG} - K * (\text{Raw EOG})$$

- Corrected EEG epochs then averaged together to get blink-corrected ERP

Validity of Ocular Correction

- Can produce valid results, but important to examine data to ascertain how well procedure worked.
- Variant of Gratton et al devised by Semlitsch, Anderer, Schuster, and Presslich (1986).
 - Creates blink-locked averages
 - Should reduce event-related contributions to correction estimate
 - Produces highly similar results



Four methods of undetermined validity for dealing with Blink Artifact in an Oddball Paradigm. Solid lines represent frequent novel items, and dotted lines represent rare learned items.

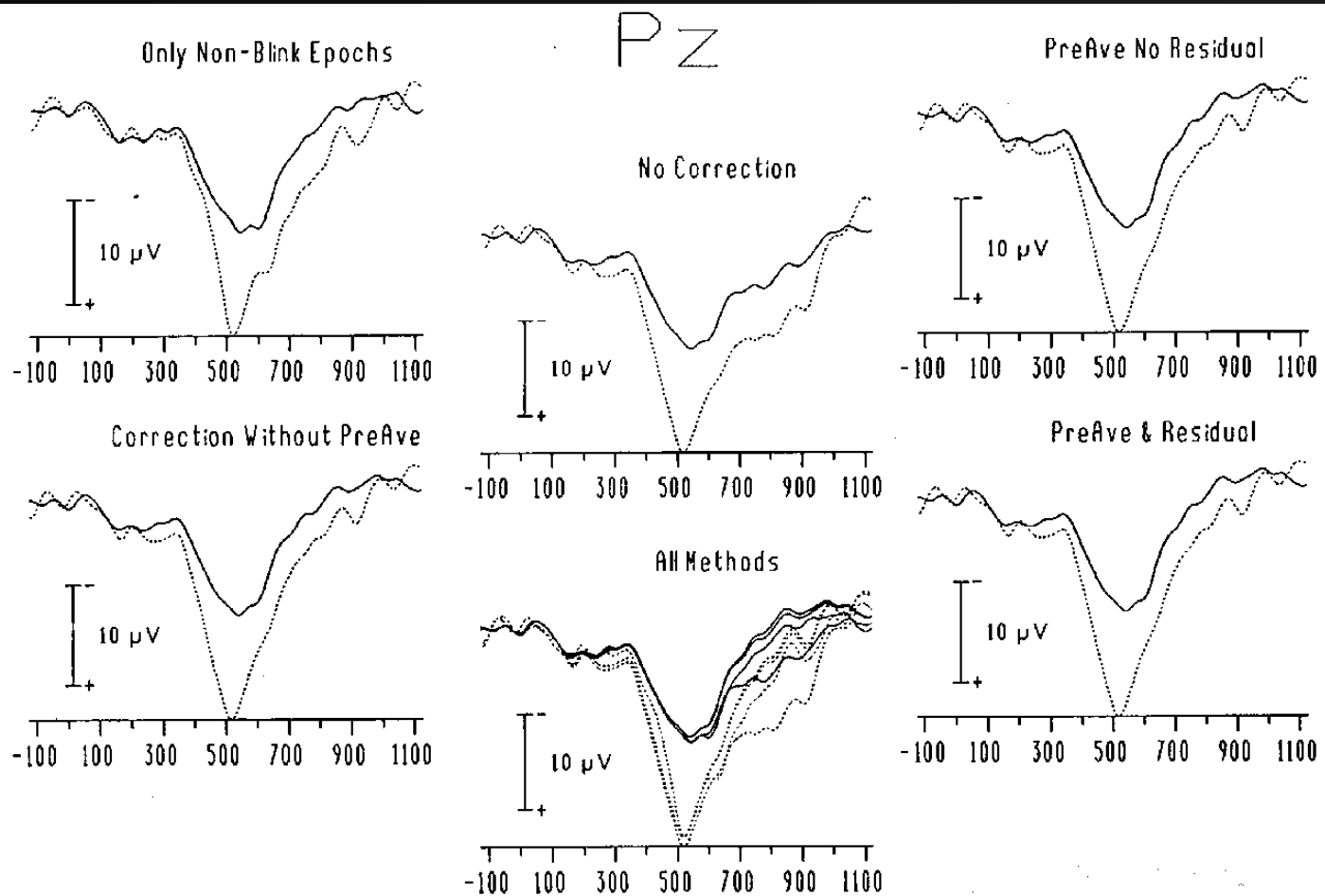
"Only Non-Blink Epochs" = excluding blink-contaminated epochs (28/60 Learned, 34/150 Unlearned)

"Correction without PreAve" = Gratton et al. method without the preliminary subtraction of event-related activity

"PreAve No Residual" = Gratton et al. method, event-related activity extracted prior to correction, no residual correction

"PreAve & Residual" = Gratton et al. method, event-related activity extracted prior to correction, with residual correction

For comparison, non-corrected data and all methods are presented in the center column. Abscissa is latency (msec).



Four methods of undetermined validity for dealing with Blink Artifact in an Oddball Paradigm. Solid lines represent frequent novel items, and dotted lines represent rare learned items.

"Only Non-Blink Epochs" = excluding blink-contaminated epochs (28/60 Learned, 34/150 Unlearned)

"Correction without PreAve" = Gratton et al. method without the preliminary subtraction of event-related activity

"PreAve No Residual" = Gratton et al. method, event-related activity extracted prior to correction, no residual correction

"PreAve & Residual" = Gratton et al. method, event-related activity extracted prior to correction, with residual correction

For comparison, non-corrected data and all methods are presented in the center column. Abscissa is latency (msec).

Other Methods (in brief)

- Most other methods also depend upon subtraction of a proportion of the EOG signal or some transformation of the EOG signal
 - Frequency-domain methods recognize that not all frequencies in the EOG channel propagate equally to scalp sites
 - Source localization methods attempt to derive a source that represents the equivalent of the origin of the eye potentials, and then compute the extent to which these sources would project onto scalp
 - BESA
 - ICA

Demonstration of Ocular Correction

One more advanced topic...

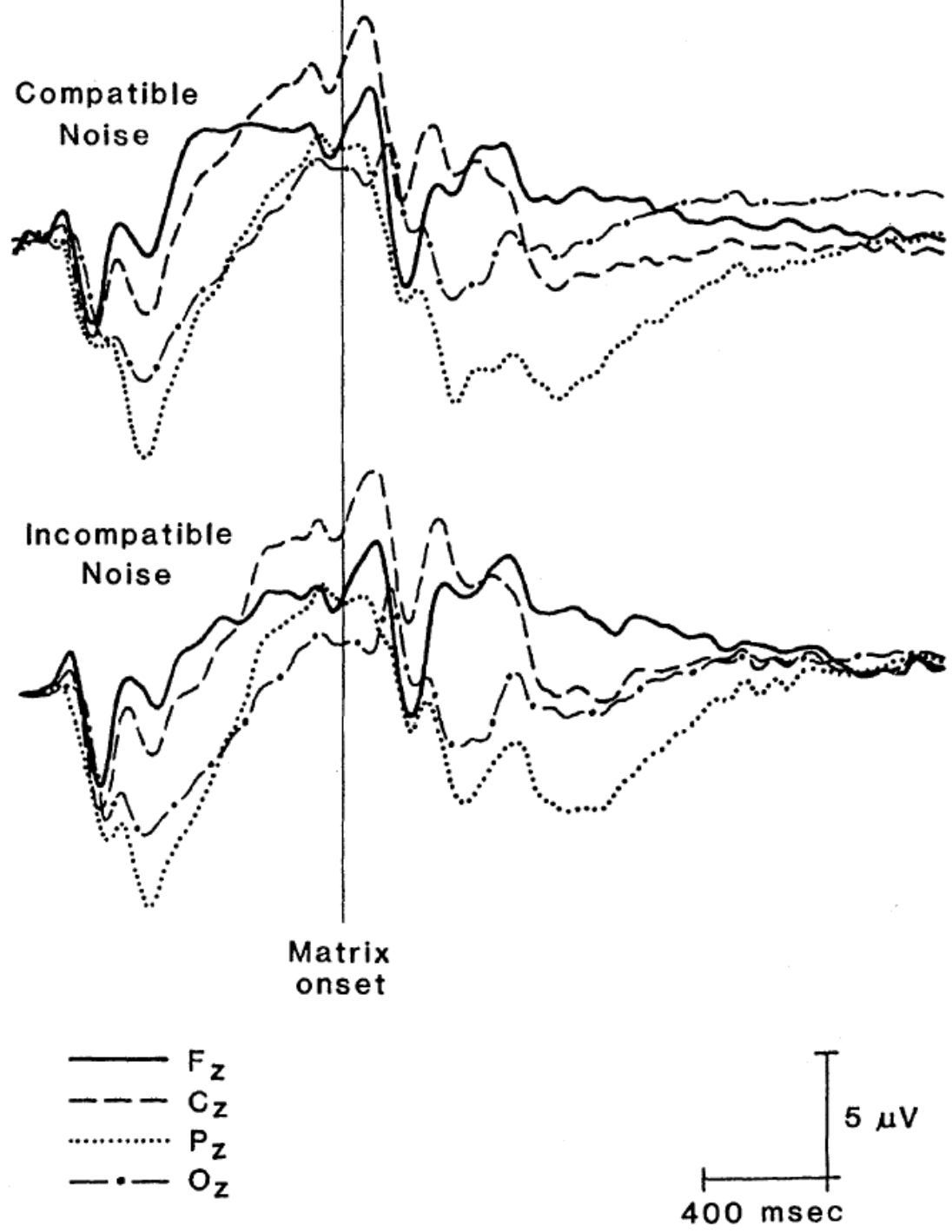
The Problem of Latency Jitter

- 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	b
Noise	
NRIGHT	KWSMNT
BMJUKM	UYRMUD
EQEIKM	VTFMZS
KEHEHG	ILEFTA
c	d

1°



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

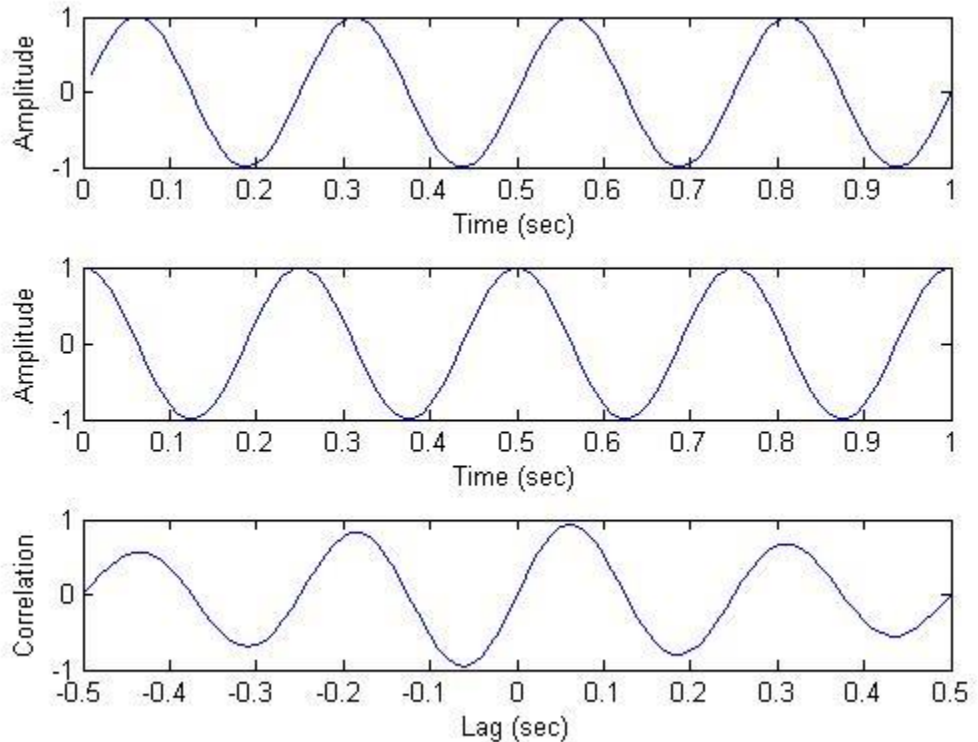
- Cross-correlational series
 - For two waveforms the correlation between each of them is computed
 - first with no lag in time
 a_1, a_2, \dots, a_n
 b_1, b_2, \dots, b_n
 - then with one lagged with respect to the other
 a_1, a_2, \dots, a_{n-1}
 b_2, b_3, \dots, b_n
 - A series of correlation values is obtained by progressively increasing the size of the lag

The Basic Idea

Sine

Cosine

Cross-
Correlation



See ... CrossCorr_Sin_Cos.m

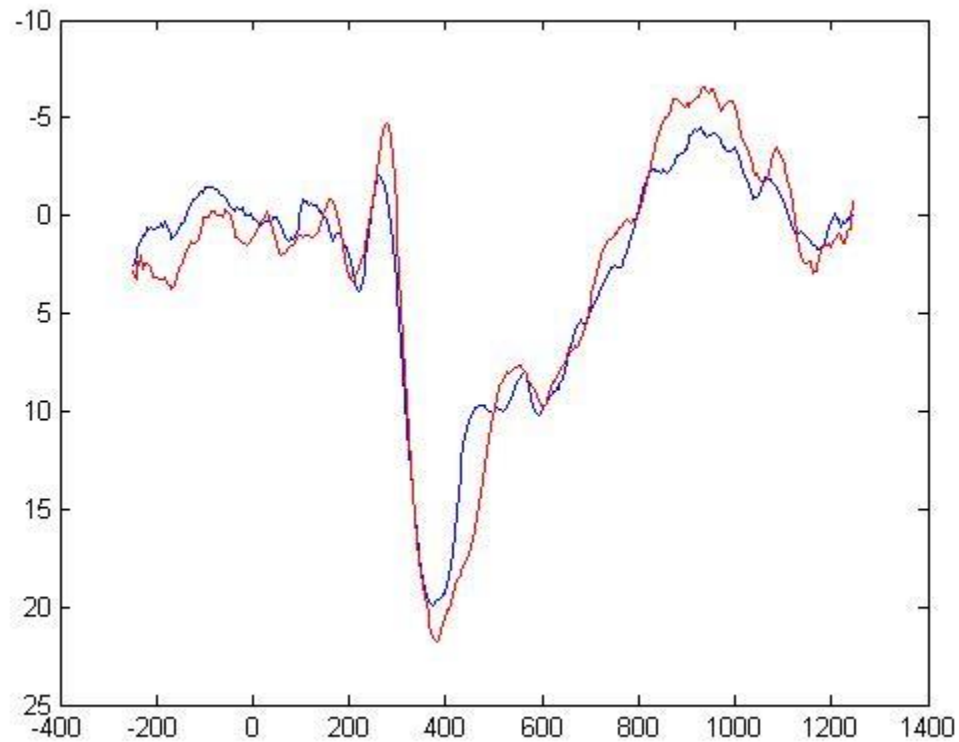
More Details

- 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-trial shifts)

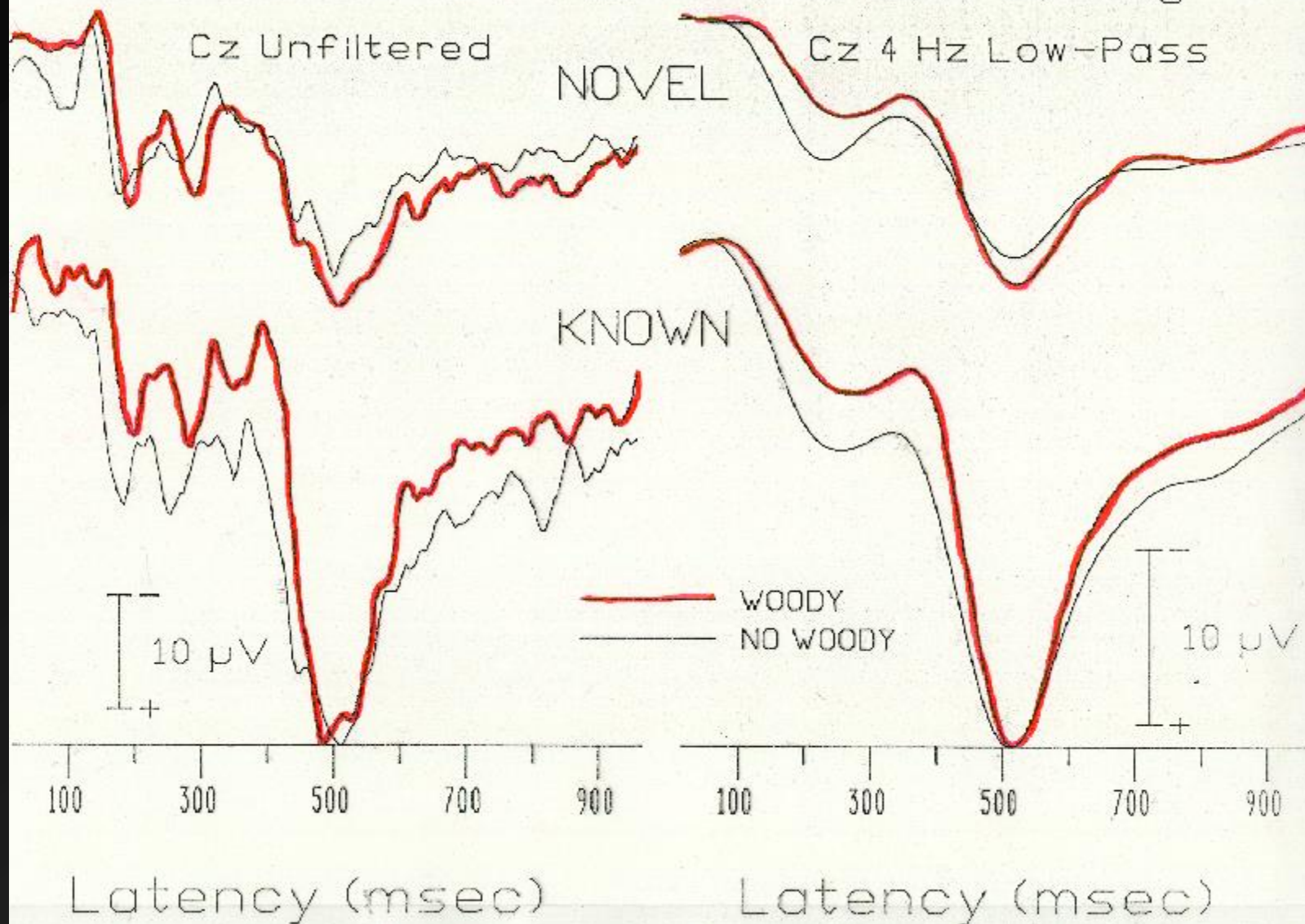
Woody's Instantiation

- 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
- Procedure
 - Initial 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
 - repeated until the maximal values of the cross correlation become stable
 - often, average cross-correlation value increment monitored; if r increases $< .005$ or $.001$, then stability achieved
- Some implementations, trials which do not reach a minimum criterion (e.g., .30-.50) are discarded from subsequent template construction and perhaps from subsequent analysis altogether

Woody Filtering Demo!



Odd-Ball ERP's sans/with WOODY Filtering



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 limit
- Suggests also that Woody may not be the solution under conditions of very low signal-to-noise ratio