THE UNIVERSITY OF RHODE ISLAND

NeuralPC Lab

Brain Signal processing & Applications

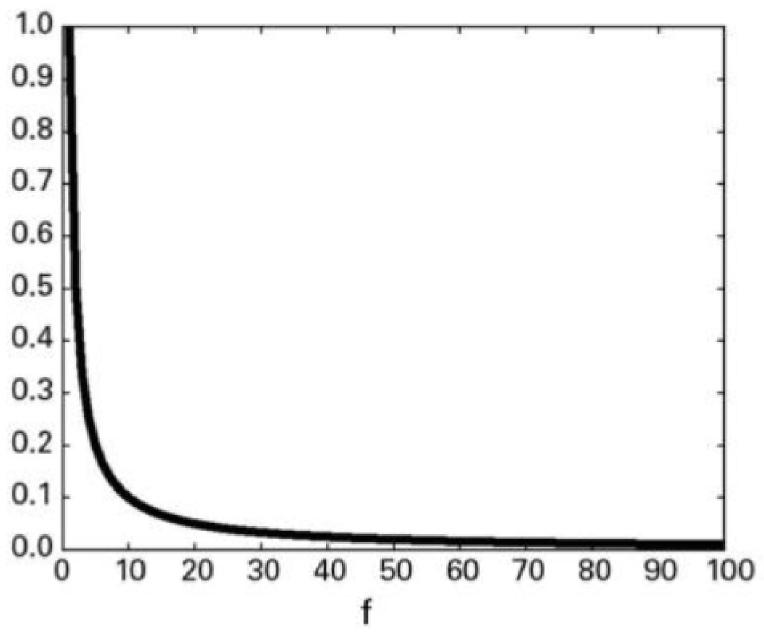
BME 473/ELE 573

Lecture 5

Dr. Yalda Shahriari

Time-Frequency Power and Baseline Normalization

- **Power law in EEG**: EEG power spectrum obeys a $1/f$ phenomenon, the power at higher frequencies (e.g., gamma) has a much smaller magnitude than the power at lower frequencies.



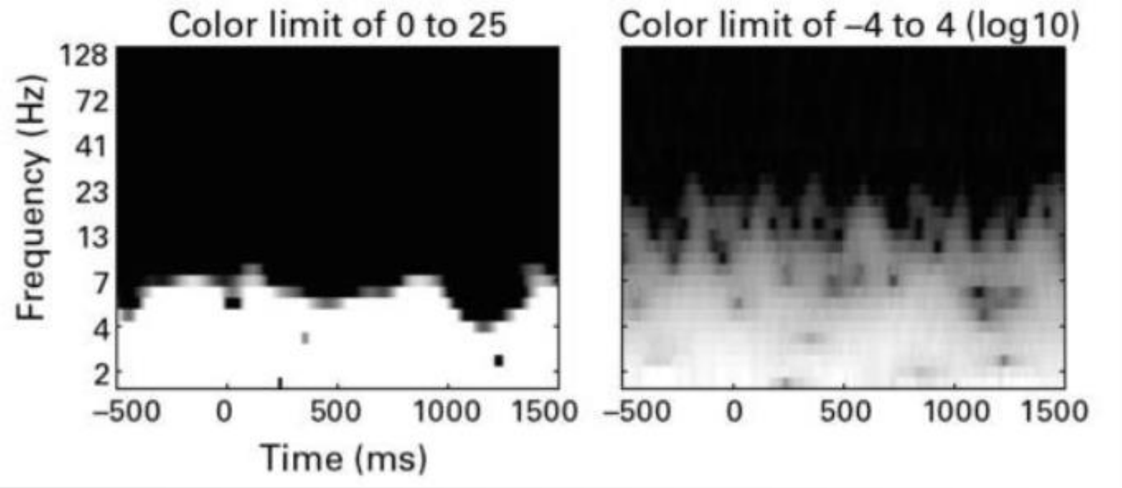
Time-Frequency Power and Baseline Normalization

Problems of raw power:

- It is difficult to visualize activity from a large range of frequency bands simultaneously.
- It is difficult to do quantitative comparison of power across frequency bands (e.g., increase of gamma,…). Thus, Interpretation gets difficult.
- Aggregating effects across the subjects might be difficult due to the influence of skull thickness, sulcal anatomy, cortical surface area,… on the raw power.
- The task related changes can be difficult to be distinguished from the background activity.
- Raw power values do not have a normal distribution (they don't have negative values, strongly positive skewed). Thus, parametric statistical analysis can not be applied to power data.

Time-Frequency Power and Baseline Normalization

$1/f$ shape can be attenuated by taking the logarithm of the power.



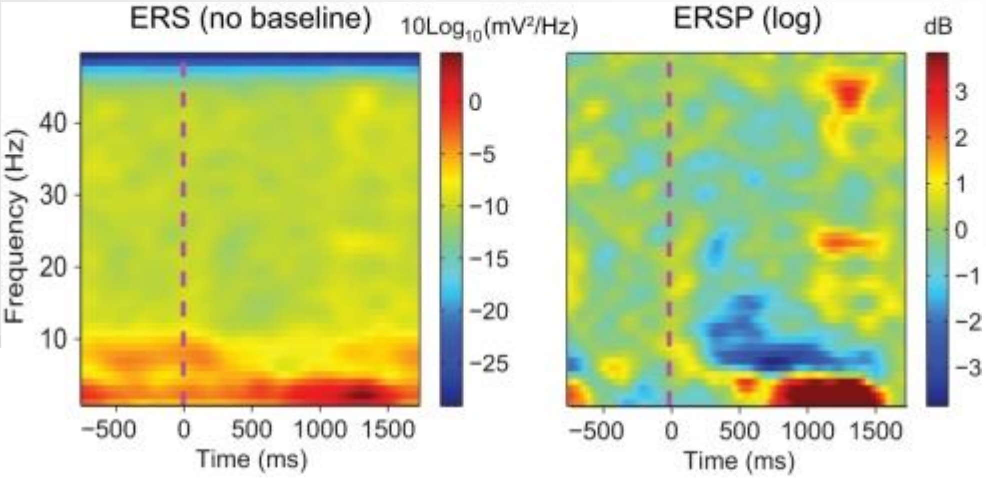
- ✓ The logarithm of the signal is more normal than the distribution of the original signal.
- ✓ Logarithmic scales allow visualizing a wider range of power variations.

But it does not completely solve our issue!

Time-Frequency Power and Baseline Normalization

- **Solutions: Baseline normalization**
- **What is a baseline?**
- ✓ A period in time (a few hundred milliseconds before the start of the trial), where no task related-activity is expected.
  - ✓ First get the average over the power of the trials and then apply baseline correction.

Difference between baseline corrected and non-baseline corrected time-frequency plots



Romain Grandchamp and Arnaud Delorme, 2011

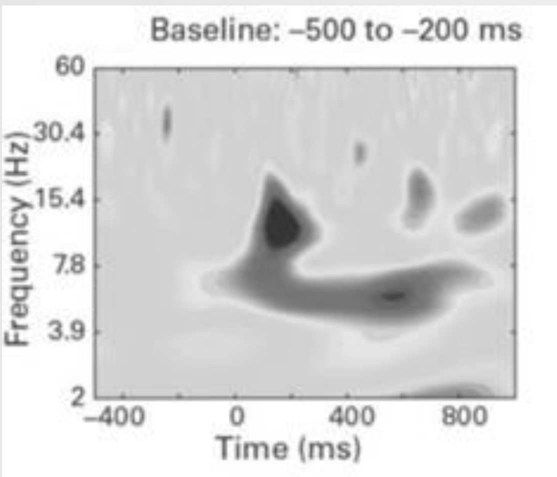
Time-Frequency Power and Baseline Normalization

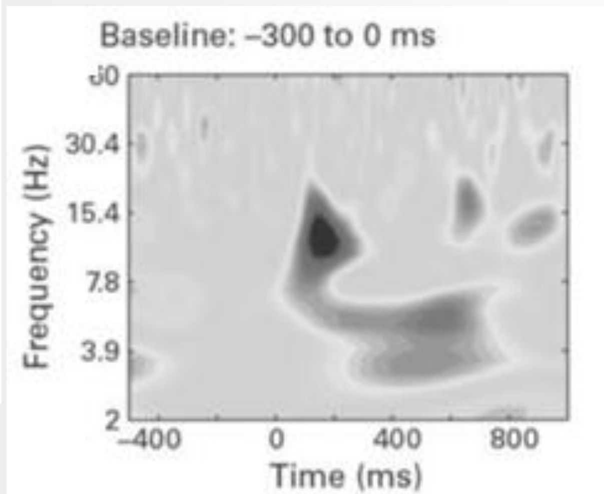
- What baseline normalization does?

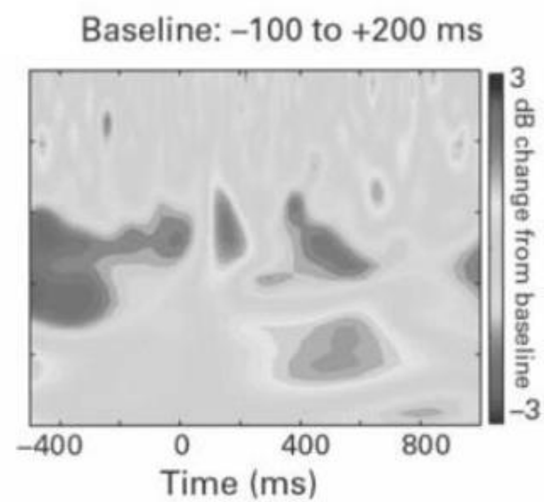
1. It transforms all power to the same scale. Thus, we can compare different frequency bands/conditions visually/statistically.
2. Task related activities become more distinguishable from background activity (task-unrelated dynamics) .
3. Makes the interpretation easier.
4. Baseline-normalized power data are often normally distributed, thus, we can apply parametric statistical tests.

Time-Frequency Power and Baseline Normalization

Difference between different choices of the baseline







Typically for the ERP analysis we can get the baseline that ends at t=0. However, for the time-frequency analysis, as we do smoothing, due to temporal leakage it's safer to end it -100 or -200 ms before the stimulus (e.g., -500 to-200ms or -400 to -100ms)

Time-Frequency Power and Baseline Normalization

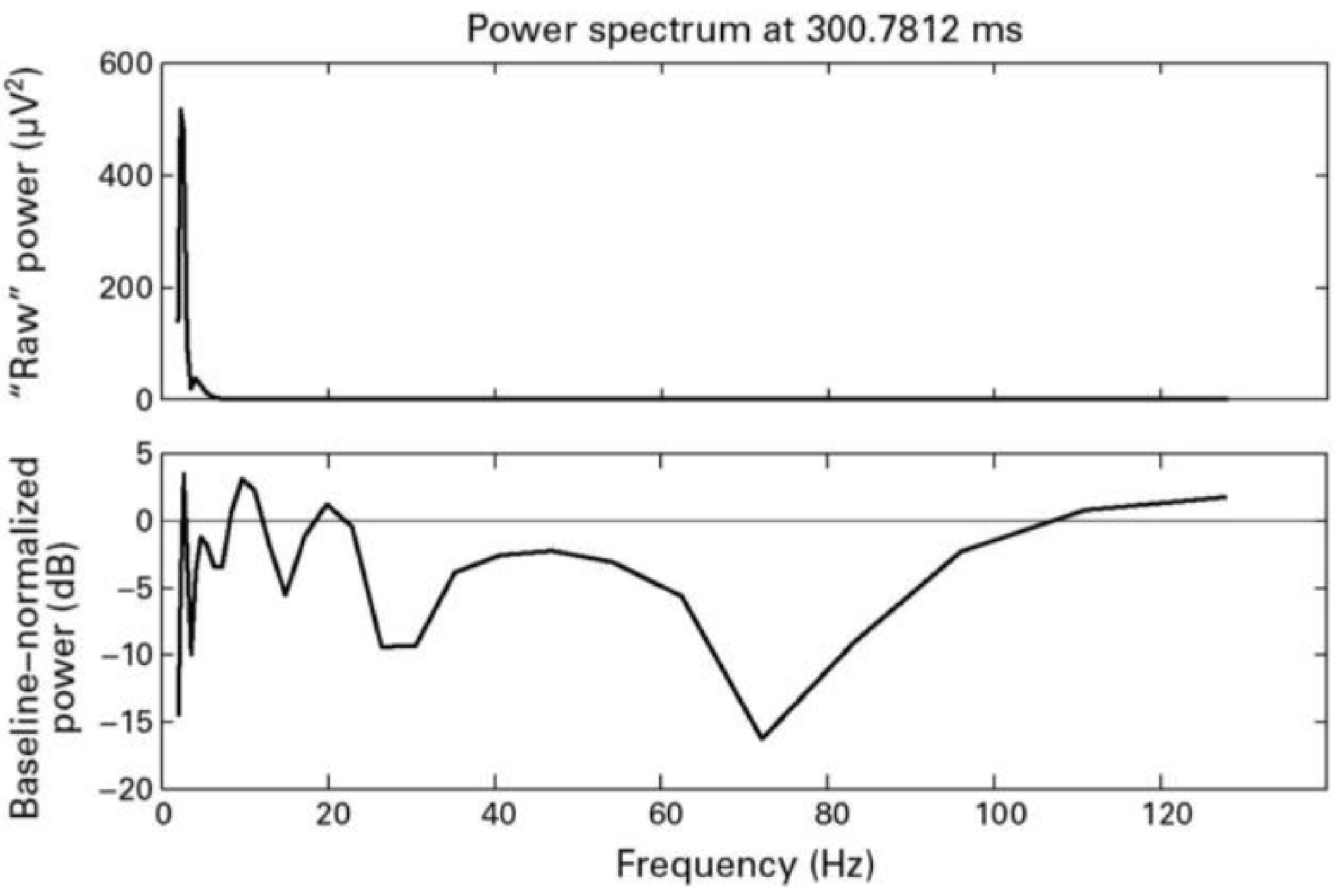
- **Three ways to do baseline normalization:**

1. The decibel (dB): is a ratio between the strength of one signal (frequency-band-specific power) and the strength of another signal (a baseline power in that same frequency band)

$$dB_{tf} = 10\log 10\left(\frac{activity_{tf}}{\overline{baseline_{f}}}\right)$$

Note: baseline does not have any “t” as we get average over all the baseline time period.

Time-Frequency Power and Baseline Normalization



Time-Frequency Power and Baseline Normalization

2. Percentage Change and Baseline Division:

$$prctchange_{tf} = 100\,\frac{activity_{tf} - \overline{baseline_f}}{\overline{baseline_f}}$$

The results are interpreted as changes in the power during baseline period

3. Z-Transform: the power data is scaled to standard deviation unites relative to the power data during the baseline period.

- ✓ The units are normal Z values and can be easily interpreted and converted to p-values (e.,. Z=1.96 corresponds to a two tailed p=0.05).

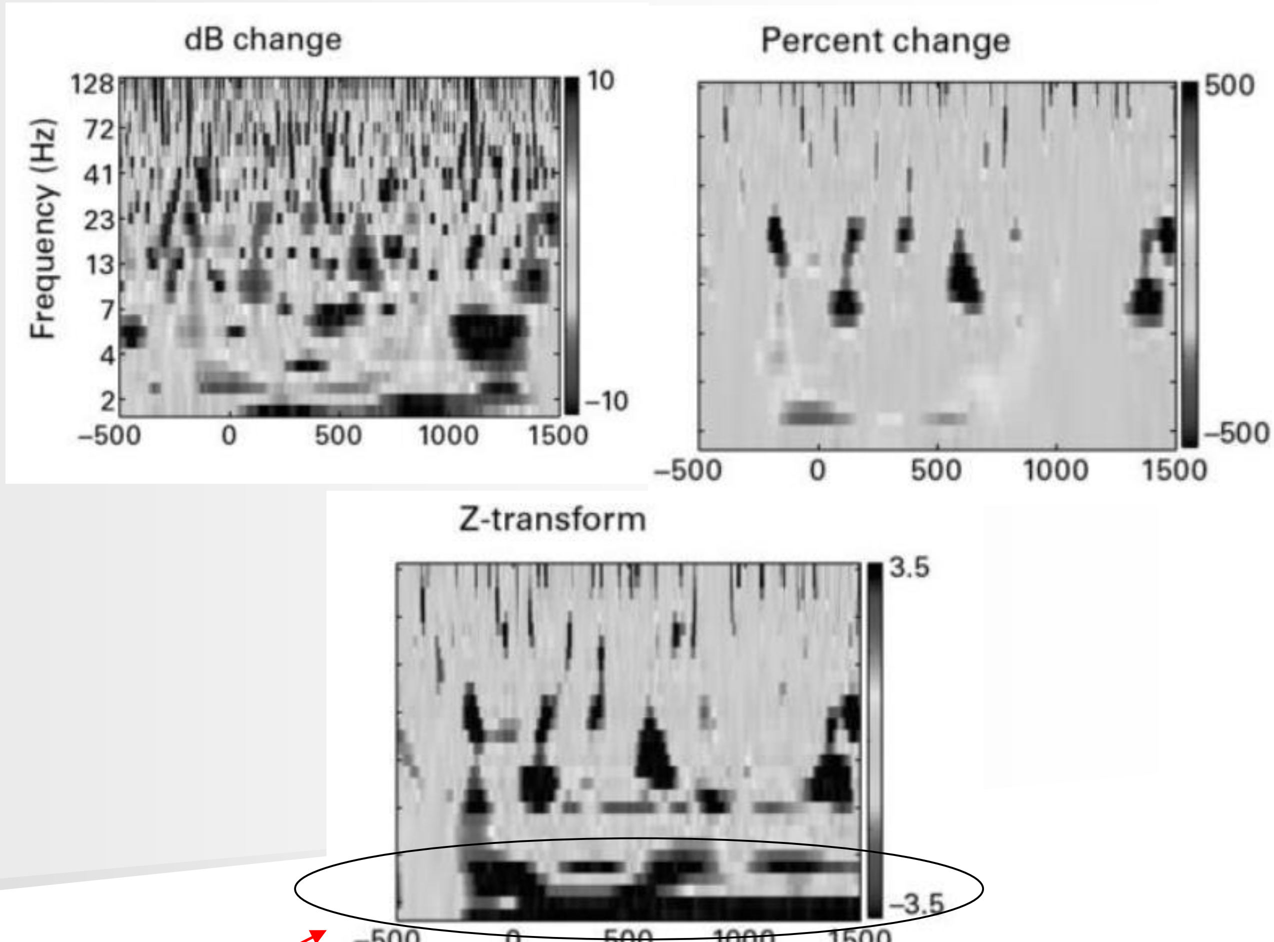
$$Z_{tf} = \frac{activity_{tf} - \overline{baseline_f}}{\sqrt{n^{-1}\sum_{i=1}^{n}\left(baseline_{if} - \overline{baseline_f}\right)^2}}$$

Standard devation of the baseline

If the data is noisy or if we have few trials, we might have very low Z-value. So, Z-value is not all the time robust.

Time-Frequency Power and Baseline Normalization

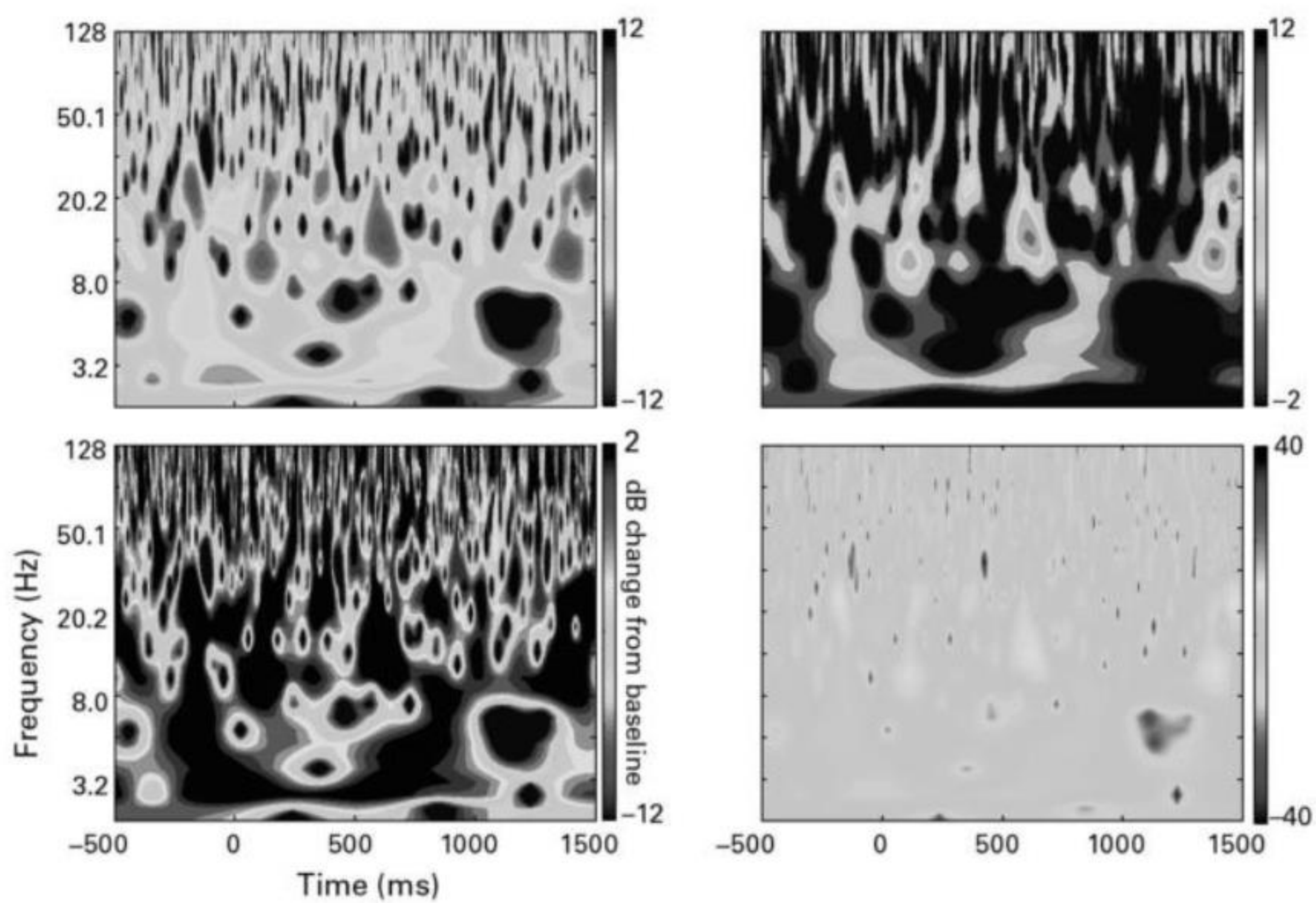
Results from the same data and same analysis, but with different baseline normalization methods applied.



Potential disadvantage of Z-transform: Large Z-values due to large std of baseline possibly due to some outliers or noise.

Time-Frequency Power and Baseline Normalization

Side Note: Effect of color scaling



Time-Frequency Power and Baseline Normalization

➢ Signal-to-Noise Estimates:

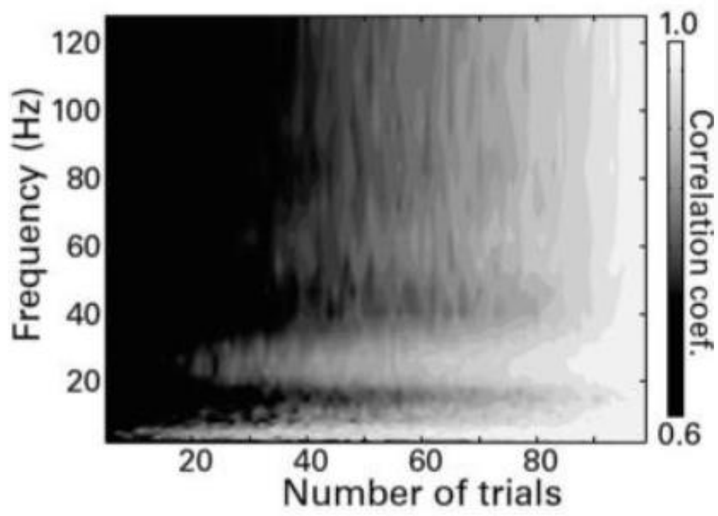
- ✓ Can be used as a measure of the quality of the data.
- ✓ Ideally we would like to dissociate signal from noise.
  - ✓ However, in real EEG they are mixed and we can not do that.
- ✓ Thus, we can estimate the SNR in an EEG data as the ratio of the mean signal to the standard deviation of the signal:

$$SNR_{tf} = \frac{\mu_{tf}}{\sigma_{tf}}$$

Where μ is the mean signal across trials and σ is the standard deviation of the signal across the trials.

Time-Frequency Power and Baseline Normalization

- ✓ One way to increase SNR, is to increase the number of trials.
- ✓ To examine how many trials is a reasonable minimum for each condition and frequency band, the average power time course of a randomly selected number of trials can be correlated with the average power time course of all trials.
- ✓ In an ideal situation with no noise, the correlation would be 1. However, the number of trials produce minimum correlation coefficient of 0.7 should be reasonable.



Hilbert Transform

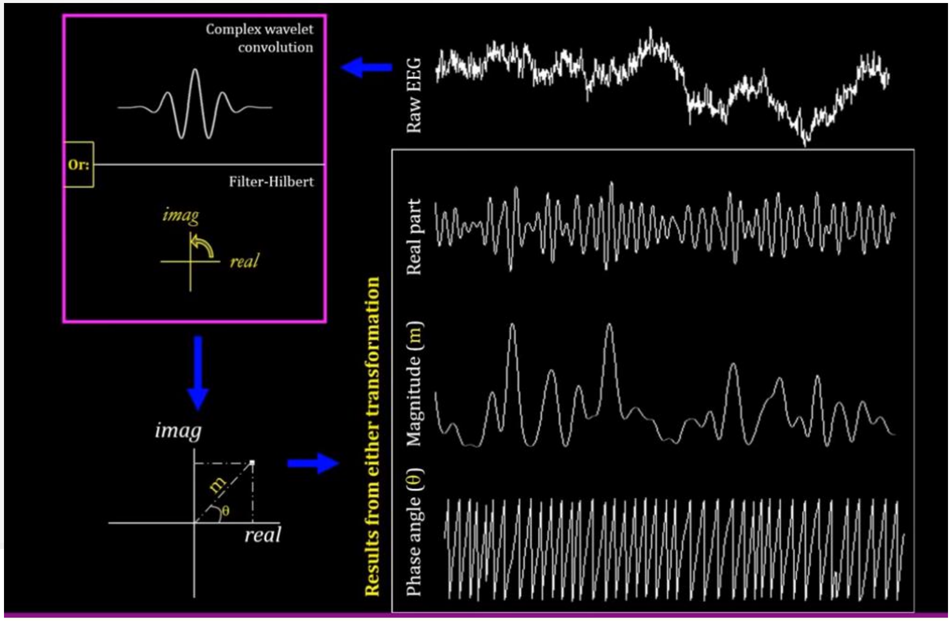
David Hilbert (1862-1943)
German Mathematician

Hilbert Transform

- Hilbert transform (HT) is another method for time-frequency decomposition: bandpass filtering and applying the Hilbert transform
  - ✓ HT allows you to extract a complex signal from a signal that contains only the real part.
  - ✓ HT is an alternative method to extract power and phase of the data.

✓ Advantage of HT vs. WT: the filter-Hilbert method allows more control over the frequency characteristics of the filter, whereas the frequency shape of a Morlet wavelet is always Gaussian.

✓ Disadvantage of HT vs. WT: bandpass filtering is a bit slower than wavelet convolution.



Hilbert Transform

- **Reminder:** To extract the power and phase from an EEG time series, we need a signal that can be represented using an Euler formula: $Me^{i2\pi ft}$ which can also be written as:

$$Me^{i2\pi ft} = M\cos(2\pi ft) + iM\sin(2\pi ft)$$

This is the analytic signal:

$$\overset{\nabla}{a}(t) \equiv a(t) + j\,\tilde{a}(t) = \left|\overset{\nabla}{a}(t)\right| \cdot e^{j\theta(t)}$$

- Hilbert transform does not involve any domain change like Laplace, Z-transform, or Fourier transform.
  - ✓ In another word, Hilbert transform of a signal x(t) is another signal denoted by $\hat{x}(t)$ in the same domain (i.e., time domain).

Hilbert Transform

- **Descriptors available through an analytic signal:**

An analytic signal: ➡

$$\overset{\nabla}{a}(t) \equiv a(t) + j\,\tilde{a}(t)$$

Magnitude

$$\left|\overset{\nabla}{a}(t)\right| = \sqrt{a^2(t) + \tilde{a}^2(t)}$$

Instantaneous Phase

$$\theta(t) = \tan^{-1}\frac{\tilde{a}(t)}{a(t)}$$

Instantaneous Frequency

$$f_i(t) = \frac{1}{2\pi}\,\frac{d\theta(t)}{dt}$$

Hilbert Transform

- The Hilbert transform of a signal is another signal whose frequency components lag the frequency component of original signal by 90°.
  - In another word, the frequency component of $\hat{x}(t)$ is exactly the same as x(t) with the same amplitude, except a 90° phase delay.

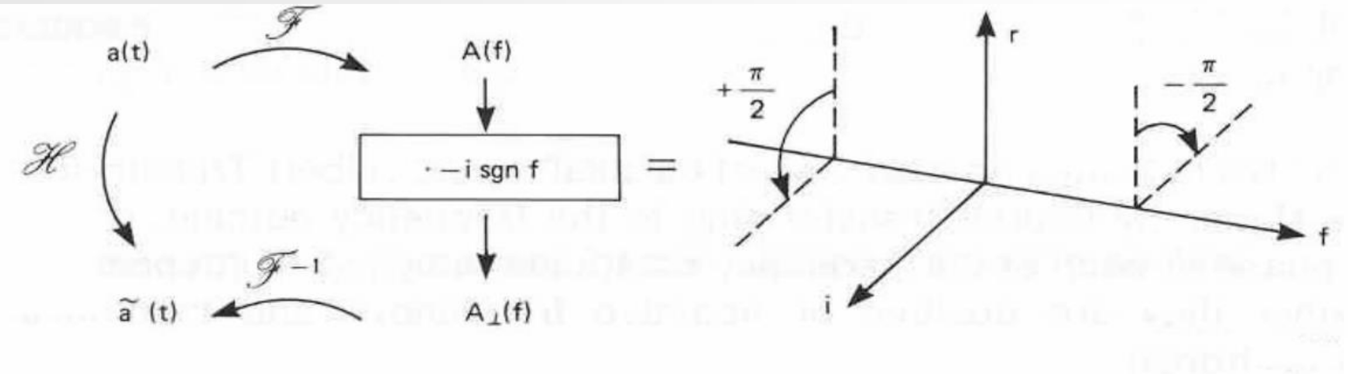
$$x(t) = A\cos(2\pi f_0 t + \theta)$$

$$\Big\downarrow H$$

$$A\cos(2\pi f_0 t + \theta - 90°) = A\sin(2\pi f_0 t + \theta)$$

Hilbert Transform

- The Hilbert transform of a signal is another signal whose frequency components lag the frequency component of original signal by 90°.
  - In another word, the frequency component of $\hat{x}(t)$ is exactly the same as x(t) with the same amplitude, except a 90° phase delay.
- **How Hilbert transform performs:**



Hilbert Transform

- **Relationship with Fourier Transform:** This is equivalent to saying that the spectrum (Fourier transform) of the signal is multiplied by -*j* sign(*f*).

$$\sigma_{\mathrm{H}}(\omega) = \begin{cases} i = e^{+\frac{i\pi}{2}}, & \text{for } \omega < 0, \\ 0, & \text{for } \omega = 0, \\ -i = e^{-\frac{i\pi}{2}}, & \text{for } \omega > 0. \end{cases}$$

- 90° (-j) for positive frequencies

+90° (j) for negative frequencies

$$F\left[\hat{x}(t)\right] = -j\,\mathrm{sgn}(f)\,X(f)$$

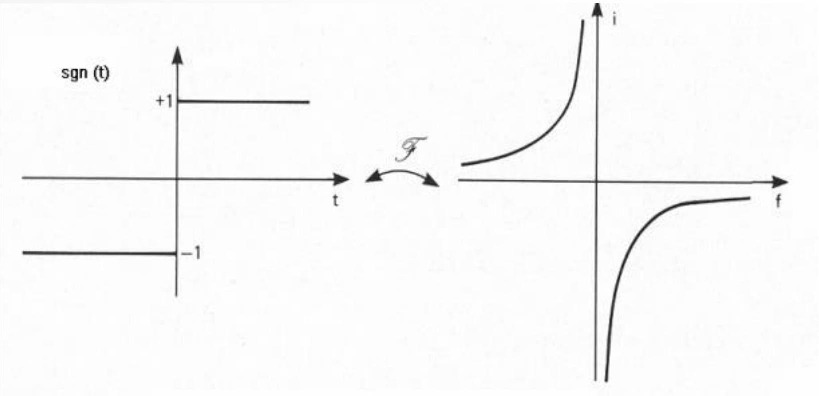
Hilbert Transform

- **How Hilbert transform performs:**

$$F\left[\hat{x}(t)\right] = -j\,\mathrm{sgn}(f)X(f)$$

$$F^{-1}\left[-j\,\mathrm{sgn}(f)\right] = \frac{1}{\pi t}$$

$$\hat{x}(t) = \frac{1}{\pi t} * x(t) = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{x(\tau)}{t-\tau}\,d\tau$$

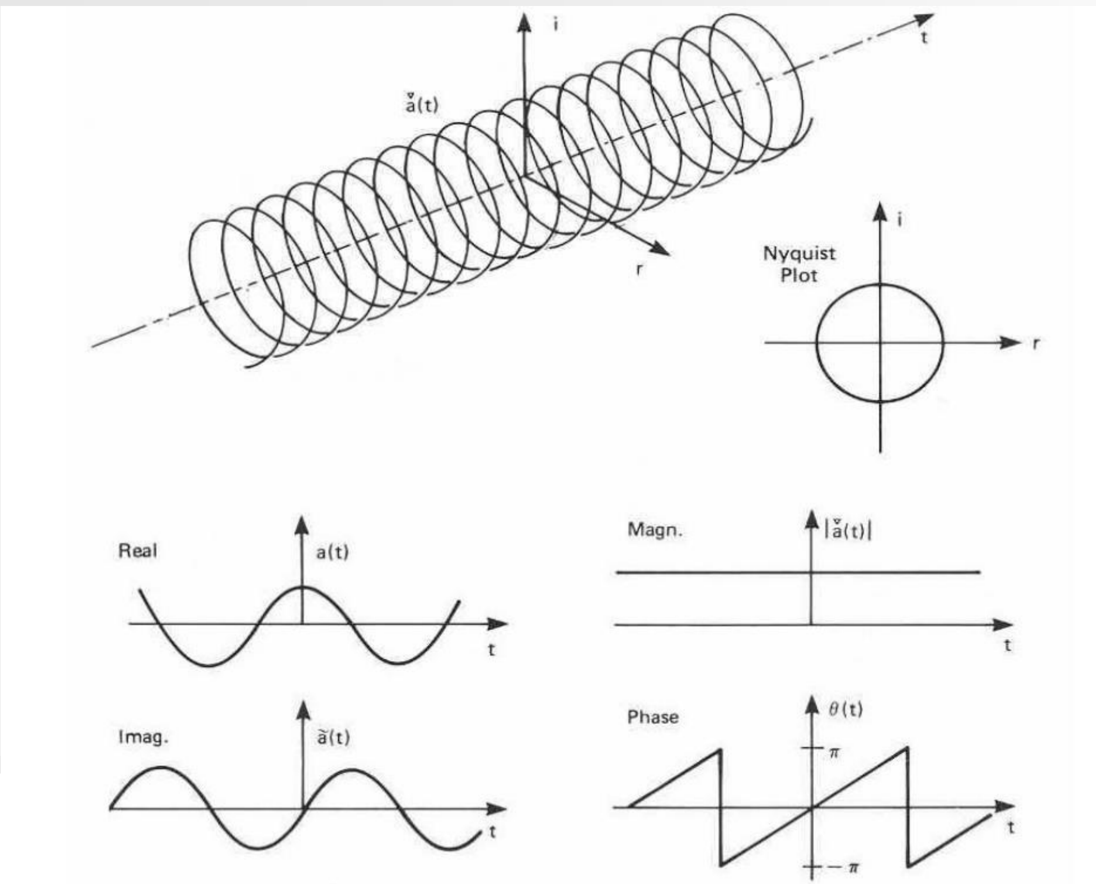


The operation of the Hilbert transform is equivalent to a convolution, i.e., filtering

Hilbert Transform

Example:

Sinusoidal Analytic Signal shown as "Heyser Spiral"



Real	$a(t) = \cos 2\pi f_0 t$
Imaginary	$\tilde{a}(t) = \sin 2\pi f_0 t$
Magnitude	$\left\lvert \overset{\nabla}{a}(t) \right\rvert = 1$
Phase	$\theta(t) = 2\pi f_0 t$
Instantaneous Frequency	$f_i(t) = f_0$

Hilbert Transform

- **Properties of Hilbert Transform:**
- ✓ **Linear Operator:**

$$\mathcal{H}\left\{ax(t)+by(t)\right\}=\mathcal{H}\left\{ax(t)\right\}+\mathcal{H}\left\{by(t)\right\}=a\mathcal{H}\left\{x(t)\right\}+b\mathcal{H}\left\{y(t)\right\}=a\widehat{x}(t)+b\widehat{y}(t)$$

- ✓ **Summary: Fourier Transform of Hilbert Transform**

$$x(t)\xleftrightarrow{\mathcal{H}}\widehat{x}(t) \qquad x(t)\xleftrightarrow{\mathcal{F}}X(\omega)$$

$$\mathcal{F}\left\{\mathcal{H}\left\{x(t)\right\}\right\}=\mathcal{F}\left\{\widehat{x}(t)\right\}=\mathcal{F}\left\{x(t)*h(t)\right\}=X(\omega)H(\omega)$$

$$H(\omega)=\mathcal{F}\left\{h(t)\right\}=\mathcal{F}\left\{\frac{1}{\pi t}\right\}$$

$$\operatorname{sgn}(t)\xleftrightarrow{\mathcal{F}}\frac{2}{j\omega} \qquad \Rightarrow \qquad \operatorname{sgn}(\omega)\xleftrightarrow{\mathcal{F}}\frac{1}{2\pi}\cdot\frac{2}{jt}$$

Reminder: $h(t)=\dfrac{1}{\pi t}\xleftrightarrow{\mathcal{F}}-j\operatorname{sgn}(\omega)=H(\omega)$

Therefore: $\widehat{X}(\omega)=-j\operatorname{sgn}(\omega)\,X(\omega)=-j\operatorname{sgn}(\omega)\mathcal{F}\left\{x(t)\right\}$

Hilbert Transform

- **Sign Reversal:** Applying Hilbert transform operation to a signal two time → sign reversal.

Consider

$$\widehat{\widehat{x}}(t) \overset{\mathcal{F}}{\longleftrightarrow} \widehat{\widehat{X}}(\omega) = (-j\mathrm{sgn}(\omega))\left(\widehat{X}(\omega)\right) = (-j\mathrm{sgn}(\omega))\,(-j\mathrm{sgn}(x)X(\omega)) = (-j\mathrm{sgn}(\omega))^2\,X(\omega)$$

$$(-j\mathrm{sgn}(\omega))^2 = (-j)^2\mathrm{sgn}^2(\omega) = -1$$

So

$$\widehat{\widehat{x}}(t) \overset{\mathcal{F}}{\longleftrightarrow} \widehat{\widehat{X}}(\omega) = -X(\omega)$$

i.e.

$$\widehat{\widehat{X}}(\omega) = -X(\omega) \qquad \Longleftrightarrow \quad \mathcal{H}^2 X(\omega) = -X(\omega)$$

Thus:

$$\hat{\hat{x}}(t) = -x(t)$$

Hilbert Transform

✓ **Energy:**

The energy content of a signal is equal to the energy content of its Hilbert transform

Using Rayleigh's theorem of the Fourier transform we have:

$$E_x = \int_{-\infty}^{\infty} |x(t)|^2 \, dt = \int_{-\infty}^{\infty} |X(f)|^2 \, df$$

$$E_{\hat{x}} = \int_{-\infty}^{\infty} |\hat{x}(t)|^2 \, dt = \int_{-\infty}^{\infty} |-j\,\mathrm{sgn}(f) X(f)|^2 \, df = \int_{-\infty}^{\infty} |X(f)|^2 \, df$$

Hilbert Transform

- ✓ **Orthogonality:**

The signal x(t) and its Hilbert transform are orthogonal

By Parseval’s Theorem of Fourier analysis we have:

$$\int_{-\infty}^{\infty} x(t)\hat{x}^*(t)dt = \int_{-\infty}^{\infty} X(f)\widehat{X}^*(f)df$$

$$\int_{-\infty}^{\infty} x(t)\hat{x}^*(t)dt = \int_{-\infty}^{\infty} X(f)[-j\,\mathrm{sgn}(f)X(f)]^*df$$

$$= -j\int_{-\infty}^{0} |X(f)|^2 df + j\int_{0}^{\infty} |X(f)|^2 df = 0$$

Hilbert Transform

✓ **Example:**

Obtain the Hilbert transform of $x(t) = cos(w_0 t)$

Reminder:

$\cos(\omega_0 t)$	$\pi[\delta(\omega-\omega_0)+\delta(\omega+\omega_0)]$
$\sin(\omega_0 t)$	$\frac{\pi}{j}[\delta(\omega-\omega_0)-\delta(\omega+\omega_0)] =$

$$j\pi[\delta(\omega+\omega_0)-\delta(\omega-\omega_0)]$$

Thus:

$$x(t) = cos(w_0 t)$$

$$X(\omega) = \pi[\delta(\omega - \omega_0) + \delta(\omega + \omega_0)].$$

Hilbert Transform

On the other hand we know:

$$\mathcal{F}[\hat{x}] = H(\omega)X(\omega) = -j\,\mathrm{sgn}(\omega)X(\omega)$$

Assuming $w_0 > 0$ then we have:

$$-j\,\mathrm{sgn}(\omega)X(\omega) = j\pi[\delta(\omega+\omega_0) - \delta(\omega-\omega_0)]$$

$$\hat{x}(t) = \widehat{\cos(\omega_0 t)} = \mathcal{F}^{-1}[-j\,\mathrm{sgn}(\omega)X(\omega)] = \sin(\omega_0 t)$$

Hilbert Transform

- **Summary:**
- Performing the Hilbert transform on a signal is equivalent to a 90° phase shift in all its frequency components.
- Therefore, **the only change that the Hilbert transform performs on a signal is changing its phase.**
- The amplitude of the frequency components of the signal do not change by performing the Hilbert-transform.
- In a simpler language, since performing the Hilbert transform changes cosines into sines, the Hilbert transform $\hat{x}(t)$ of a signal x(t) is orthogonal to x(t).
- Also, since the Hilbert transform introduces a 90° phase shift, carrying it out twice causes a 180° phase shift, which can cause a sign reversal of the original signal.

Hilbert Transform

- **How does Hilbert transform work in Matlab works:**
- ✓ "hilbert" Matlab function basically gives the analytic signal where the real part is the original signal and the imaginary part is the Hilbert transform.

- **Step by step instruction:**

1- First compute the Fourier transform.

2- Create a copy of Fourier transform that has been multiplied by the complex operator "i".

3- Identify the positive and negative frequencies

- ✓ The positive frequencies are those between but not including the zero and the Nyquist frequencies (zero and Nyquist frequency will be left untouched)
- ✓ The negative frequencies are those above the Nyquist frequency

Hilbert Transform

4- Rotate the positive frequency Fourier coefficients one-quarter cycle counterclockwise (i.e., -90° or multiply by complex operator "-i") and the negative frequency Fourier coefficients one-quarter cycle clockwise (i.e., 90° or multiply by complex operator "i").

5- Add the rotated positive frequency Fourier coefficients to the positive frequency coefficients (the real part) and do the same for the negative frequencies.

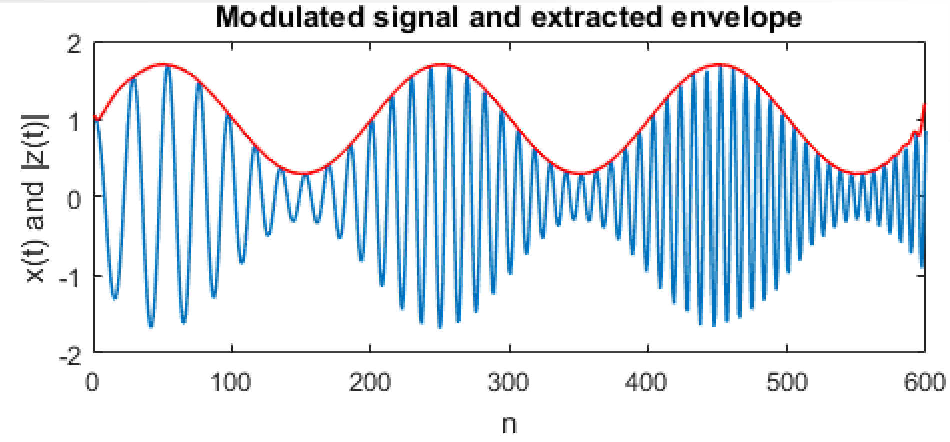
- ✓ This can be equivalent to doubling the positive-frequency coefficients and zeroing the negative-frequency coefficients.

6- Take the inverse Fourier transform.

- ❖ The results is the analytic signal which can be used in the same way as we did for Morelet wavelet. Thus we can have power, phase, and real components

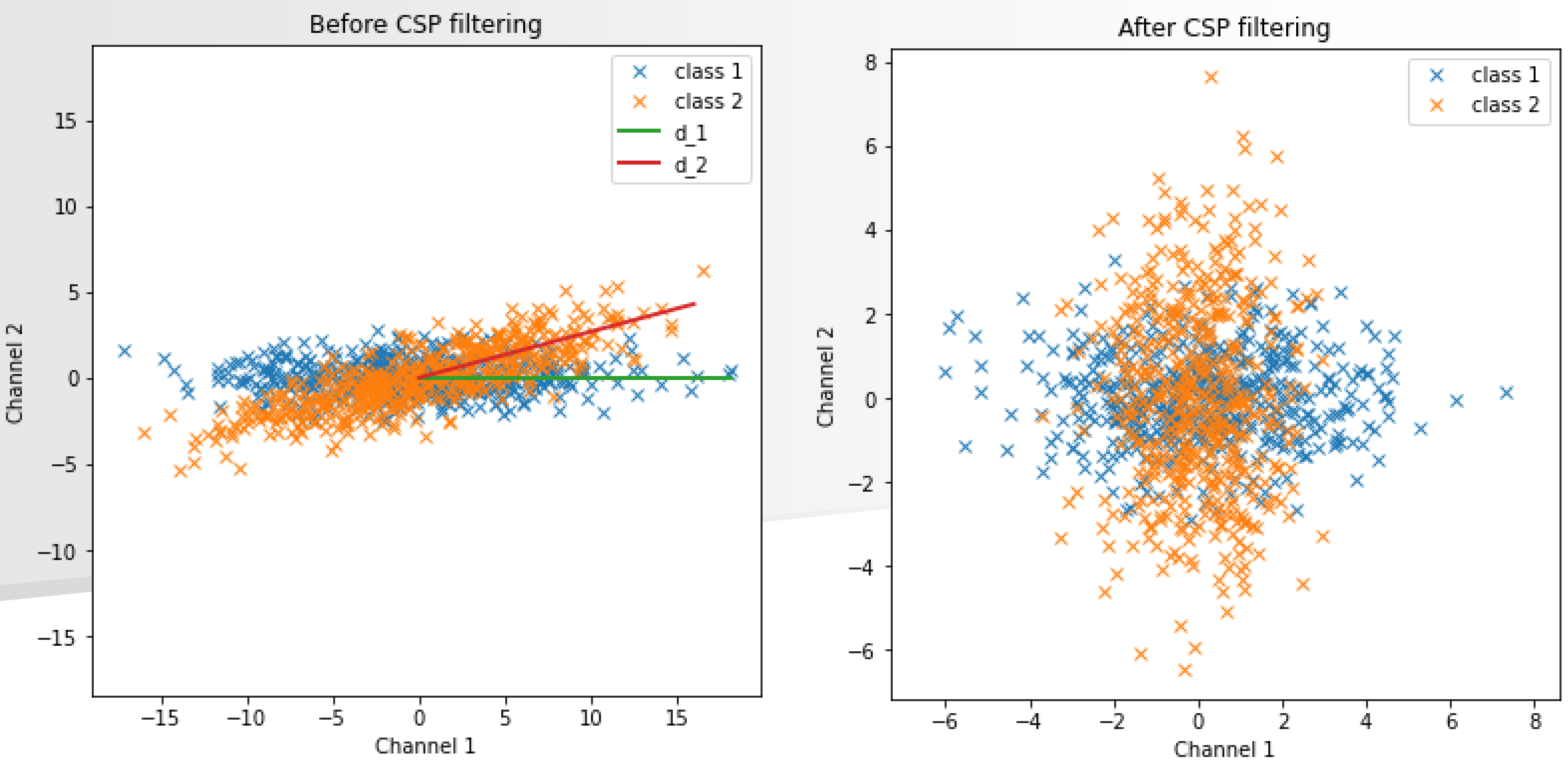
Hilbert Transform

- **The main application of Hilbert in brain signal analysis is envelope extraction.**
- ✓ To do so, filter the data into separate frequency bands before applying the Hilbert. Thus we can have, frequency-band-specific interpretation.
- ✓ Apply Hilbert transform, thus we will have the analytic signal.
- ✓ Then get the absolute value (abs) which gives the envelope.
  - ✓ Thus envelope is the abs of the analytic signal.

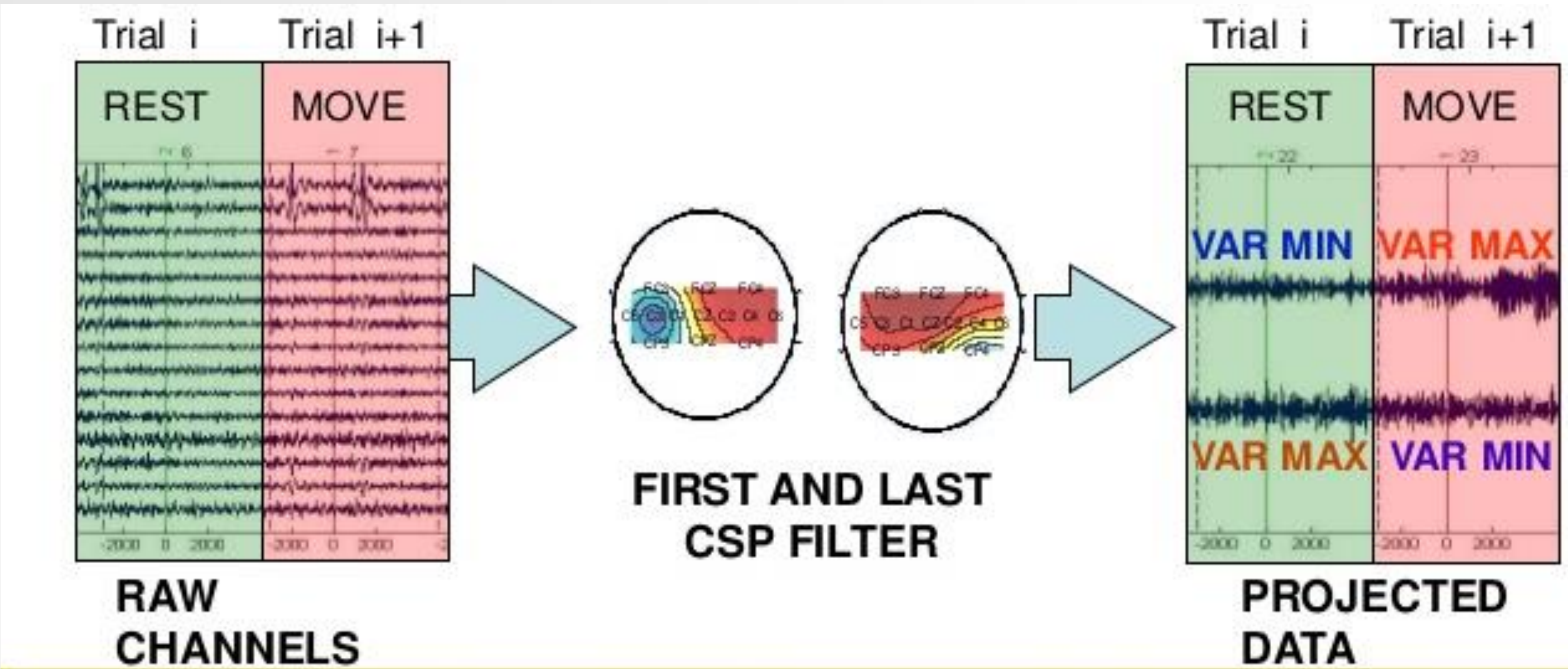


Common Spatial Pattern

- Common spatial pattern (CSP) is a supervised spatial filtering that discriminates the data from two classes by maximizing the variance of one class and minimizing the variance of the other class (Muller et al., 1999).
- **Why use CSP?** We want to make the classifier's job easier



Common Spatial Pattern



Common Spatial Pattern

Problem statement:

This is a generalized eigenvalue problem!

$$w_i = \max_{w} \; w^T \Sigma^{(i)} w \quad \text{s.t.} \quad w^T \left(\Sigma^{(1)} + \Sigma^{(2)}\right) w = 1$$

Where $\boldsymbol{w}^T \boldsymbol{\Sigma} \boldsymbol{w}$ is the variance of the projected data (variance in direction w) and w_i is the spatial filter w.

- ✓ $\Sigma^{(1)} = < \Sigma_n^{(1)} >$trials (average of covariances across all the trials for class (1))
- ✓ $\Sigma^{(2)} = < \Sigma_n^{(2)} >$trials (average of covariances over all the trials for class (2))
- ✓ $\Sigma^{(c)} = \Sigma^{(1)} + \Sigma^{(2)}$ (composite covariance matrix)
- ✓ Each $\Sigma_n^{(1)}$ or $\Sigma_n^{(2)}$ is the normalized covariance matrices for trial n of class (1) or (2).

$$\boldsymbol{\Sigma}_1 = \underset{\boldsymbol{E}_n \in \{\text{class } 1\}}{\text{Exp}} \frac{\boldsymbol{E}_n \boldsymbol{E}_n^T}{\text{tr} \boldsymbol{E}_n \boldsymbol{E}_n^T}$$

$$\boldsymbol{\Sigma}_2 = \underset{\boldsymbol{E}_n \in \{\text{class } 2\}}{\text{Exp}} \frac{\boldsymbol{E}_n \boldsymbol{E}_n^T}{\text{tr} \boldsymbol{E}_n \boldsymbol{E}_n^T}$$

Common Spatial Pattern

CSP Algorithm:

1- Covariance Matrices Calculation:

- Let E_n be the multi-channel data segment in trial n. E_n is an N x T matrix where N is the number of electrodes and T is the samples in time.

Note: Each trial should be centered (mean removed prior to the analysis).

- Compute the covariance matrix for each E_n in each class as $\sum_n^{(1)}$ and $\sum_n^{(2)}$.
- Then calculate the average covariance matrices for both classes: $\sum^{(1)} =< \sum_n^{(1)} >$ & $\sum^{(2)} =< \sum_n^{(2)} >$.

2- Composite Covariance Matrix and Whitening:

- The composite covariance matrix is $\sum^{(c)} = \sum^{(1)} + \sum^{(2)}$.
- Perform the eigenvalue decomposition on $\sum^{(c)}$

Thus, we have: $\sum^{(c)} = UDU^T$

where D is the diagonal matrix of eigenvalues, and U contains the corresponding eigenvectors.

Common Spatial Pattern

- Perform whitening on the composite covariance matrix.
  - ✓ Whitening is a process to transform the signals so that they have a unit variance and zero covariance. It simplifies the problem by removing correlations between channels.
- The whitening transformation matrix will be $P := \sqrt{D^{-1}}U^T$

3- Simultaneous Diagonalization:

- Transform the composite covariance matrices using the whitened transformation which is equal to transforming each of the individual covariances by P

$$\widehat{\Sigma}_1 = P\Sigma_1 P^T$$
$$\widehat{\Sigma}_2 = P\Sigma_2 P^T$$

✓ What this transformation accomplishes is to scale and rotate the covariance matrices of both classes such that they can be more easily compared and processed. In the whitened space, the largest differences in variance between the two classes will be more pronounced, which is exactly what the CSP algorithm aims to exploit for feature extraction.

The above equations mean that both $\widehat{\Sigma_1}$ and $\widehat{\Sigma_2}$ share the same eigenvector.

Common Spatial Pattern

- With this observation if we perform the eigendecomposition of $\widehat{\sum_1}$ we will have:

$$\widehat{\sum_1} = V\Lambda_1 V^T$$

Then we will have the same eigenvector for $\widehat{\sum_2}$ as below:

$$\widehat{\sum_2} = V\Lambda_2 V^T$$

Where Λ_1 and Λ_2 are the eigenvalues of $\widehat{\sum_1}$ and $\widehat{\sum_2}$ respectively.

✓ As a side note we know that based on our condition we had $\widehat{\sum_1} + \widehat{\sum_2} = \mathrm{I}$. Thus, any orthonormal matrices V satisfy $V^T(\widehat{\sum_1} + \widehat{\sum_2})V = \mathrm{I}$.

Thus, the summation of two eigenvalues also goes to I:

$$\Lambda_1 + \Lambda_2 = \mathrm{I}$$

This gives us the eigenvectors V, which is **the generalized eigenvalue decomposition of $\widehat{\sum_1}$ and $\widehat{\sum_2}$**, and are optimal in the least square sense for discriminating between the two populations of the whitened EEG epochs.

Common Spatial Pattern

4- Compute CSP filters as below:

$$W = P^T V$$

- ✓ These filters will project your data onto a space where the variance is maximized for one class and minimized for the other.

In the other words when we project the original data into the CSP subspace, the projected covariances would be as below:

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{ch}.$$

$$W^T \Sigma_1 W = \Lambda = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_{ch} \end{pmatrix},$$

$$W^T \Sigma_2 W = I - \Lambda = \begin{pmatrix} 1-\lambda_1 & & \\ & \ddots & \\ & & 1-\lambda_{ch} \end{pmatrix},$$

Where

- ✓ The columns of W are the spatial filters and rows of $W^{-1}$are the spatial patterns (CSPs).
- ✓ Therefore, the first CSP filter w_1 provides maximum variance of class 1, and the last CSP filter w_{ch} provides the maximum variance of class 2.

Common Spatial Pattern

5- Select the first and last m filters as below:

$$\mathbf{W}_{csp} = (\boldsymbol{w}_1 \quad \cdots \quad \boldsymbol{w}_m \quad \boldsymbol{w}_{ch-m+1} \quad \cdots \quad \boldsymbol{w}_{ch}) \in R^{ch \times 2m}$$

Thus, a filtered trial can be obtained as below:

$$S_n = W_{csp}^T E_n$$

Where S_n is a 2m x T dimension matrix.

- ✓ The first and last m rows of S_n corresponds to the projected data with maximum variance for class 1 and maximum variance for class 2 respectively.
- ✓ Typically we use 6 filters (i.e., 3 pairs), corresponding to the 3 largest and 3 lowest eigenvalues

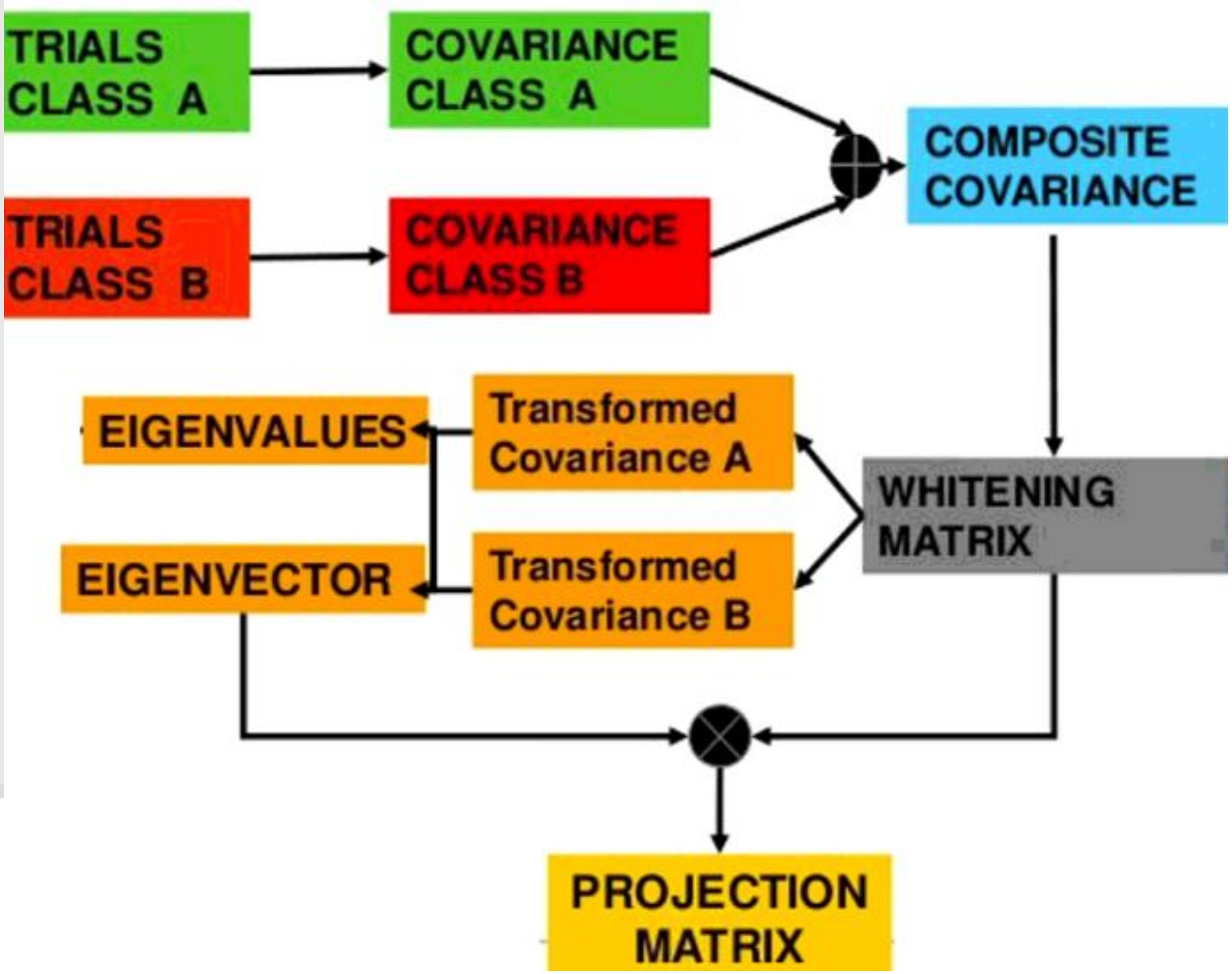
What are the feature vectors for our classifier now?!

$$x_i = \log\left(\frac{\mathrm{var}[s_i(t)]}{\sum_{i=1}^{d} \mathrm{var}[s_i(t)]}\right)$$

Note: The transformation of log is done in order to make the distributions of elements in x_i normal.

Common Spatial Pattern

- **Summary of CSP:**



Common Spatial Pattern

- **CSP Advantages:**

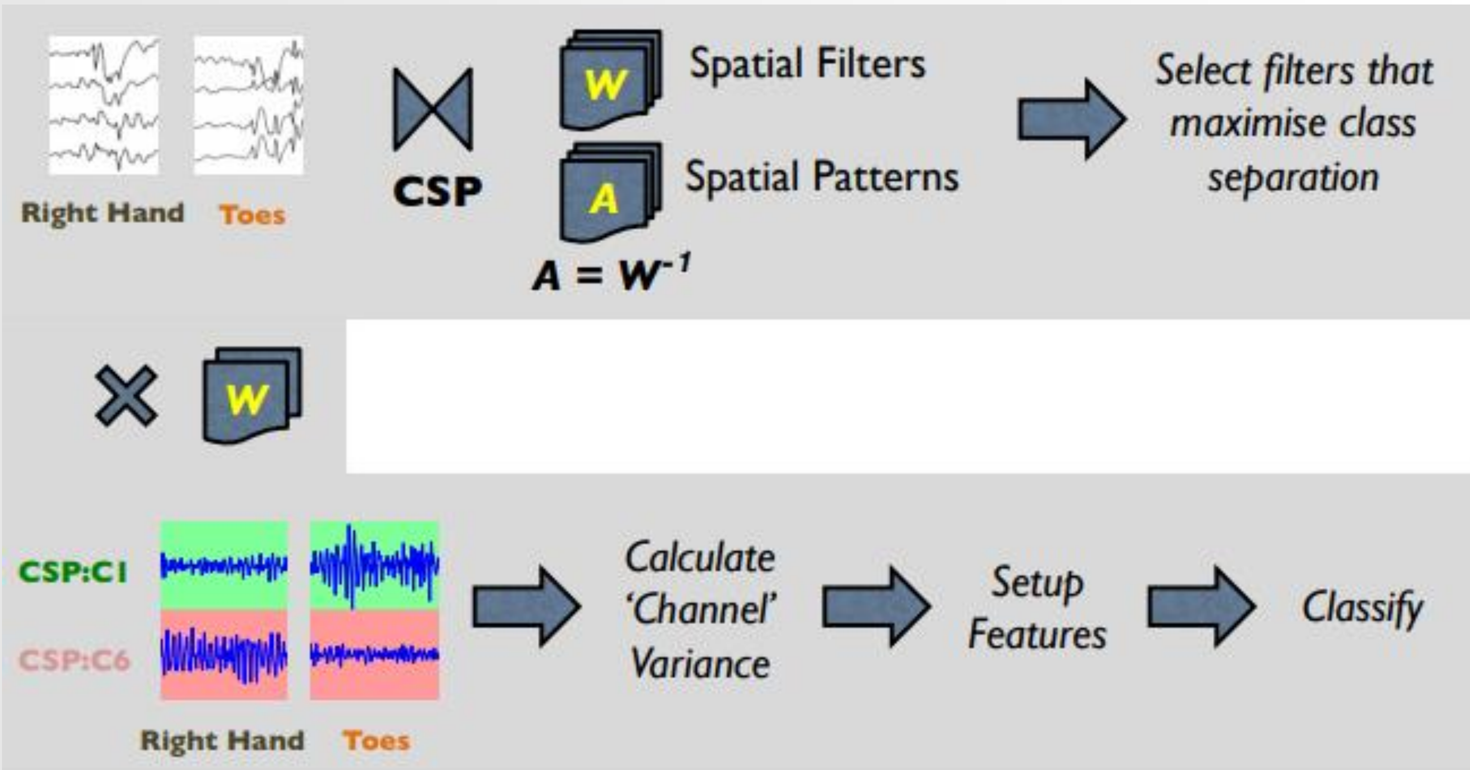
✓ CSPs can be used to give better physiological interpretation of the data.

✓ CSP filters can be used to discriminate the mental states characterized by spectral features such as ERD, MI based BCIs, etc.).

  ✓ So, we do need to band pass filter the signal prior to the analysis.

✓ Thus the obtained CSP features (based on variance) will provide the CSP features in only that specific frequency band.

Common Spatial Pattern



Apply spatial filters and test the classifier on test trial set

Common Spatial Pattern

Example:

Right hand/left hand movement and the CSP features.

- ✓ The variance of left hand and right hand has been maximized and minimized respectively in CSP1.
- ✓ The variance of right hand and left hand has been maximized and minimized respectively in CSP6