EEG/MEG 3:

Time-Frequency and Functional Connectivity Analysis

Olaf Hauk

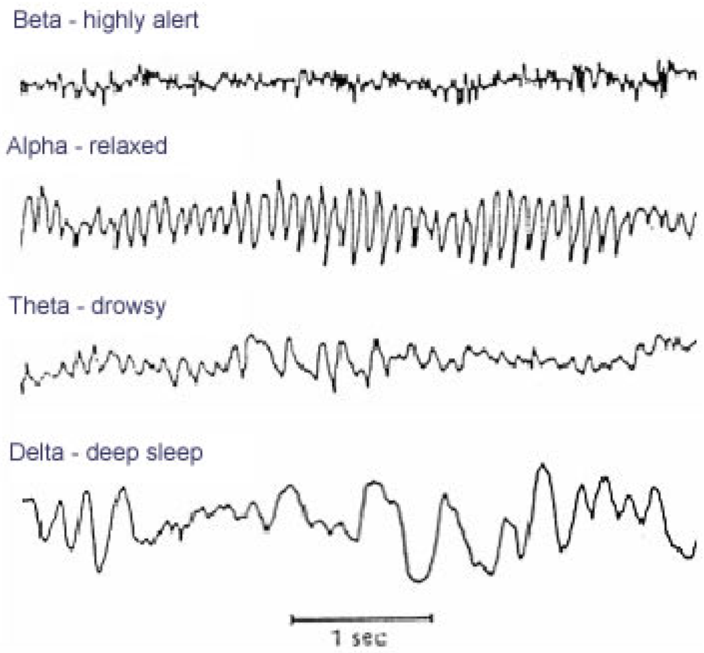
olaf.hauk@mrc-cbu.cam.ac.uk

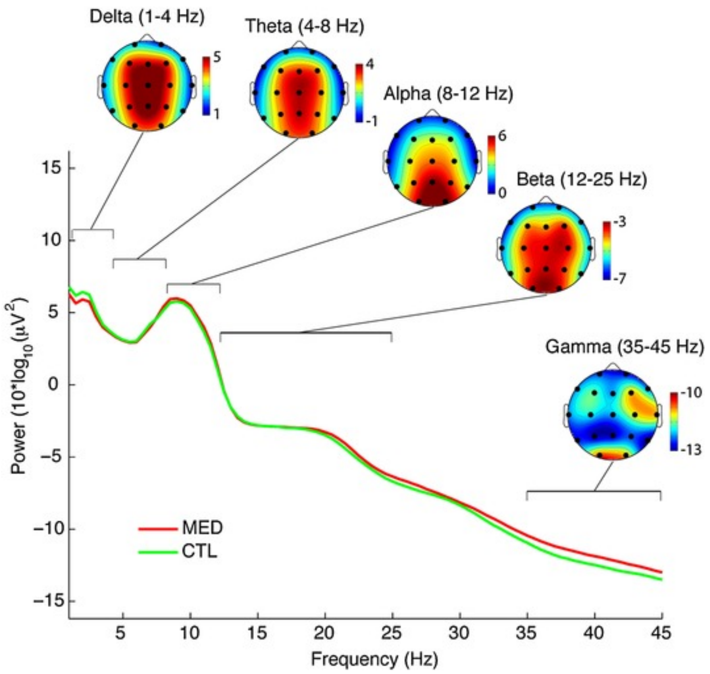
Introduction to Neuroimaging Methods, 25.1.2022

“Brain Rhythms” and “Oscillations”

Time course and topography may differ among different frequency bands

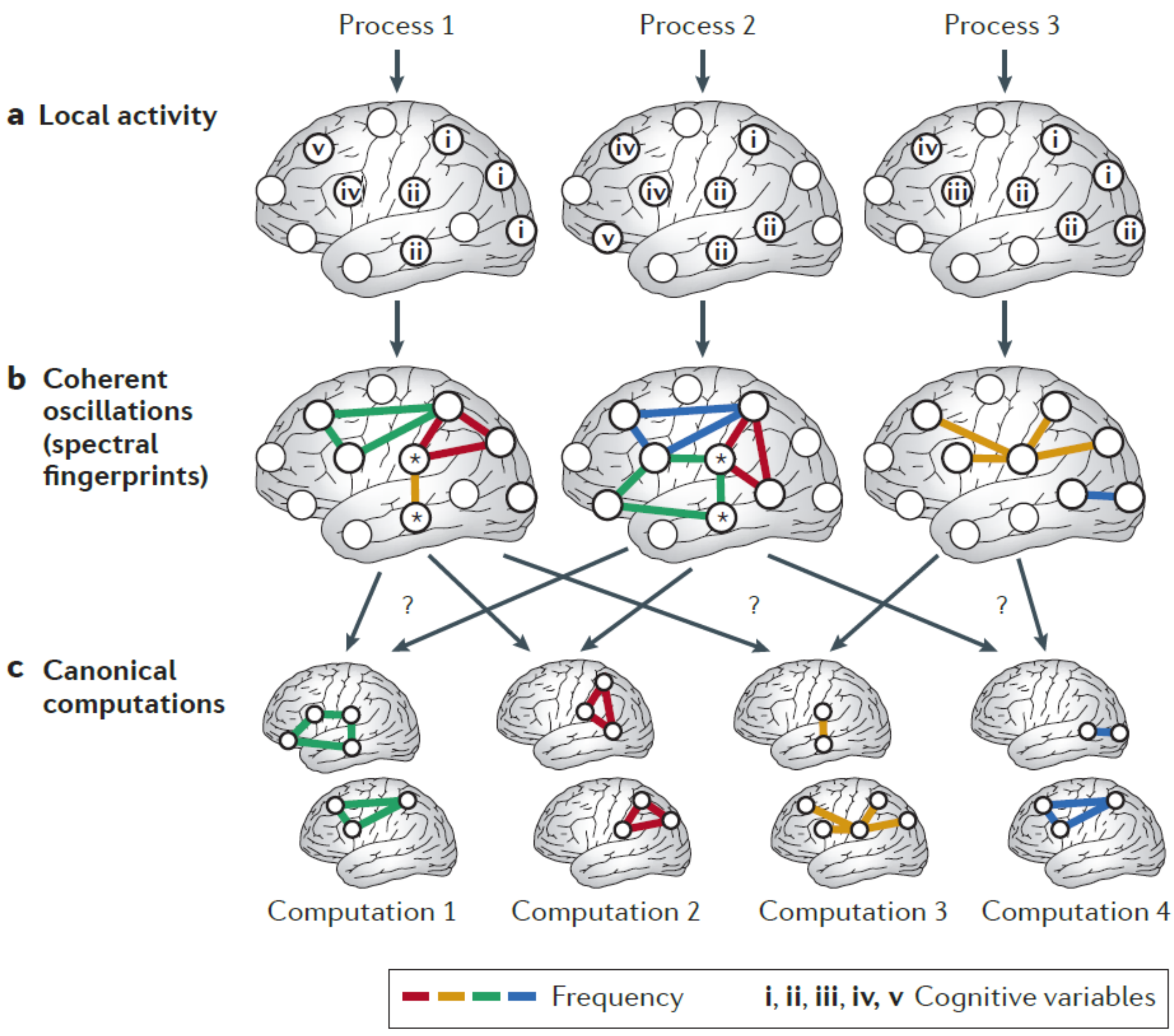
(and may depend on task, environment, subject group etc.)





Cahn et al., Cogn Proc 2010, http://link.springer.com/article/10.1007%2Fs10339-009-0352-1/

“Brain Rhythms” and “Oscillations”



Siegel et al., Nat Nsc 2012

Periodic Signals

A periodic signal repeats itself with a period T.

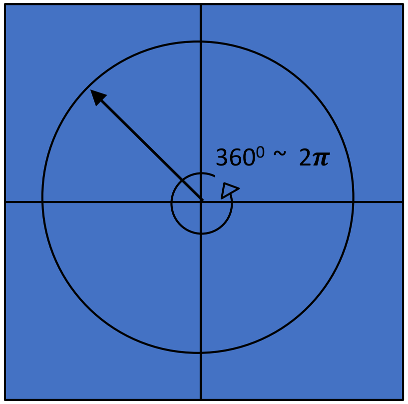
This is the case, for example, for sine and cosine functions:

In radians ($2\pi \sim 360$ degrees):

$$\cos(x+2\pi) = \cos(x)$$
$$\sin(x+2\pi) = \sin(x)$$

In degrees :

$$\cos(x+360) = \cos(x)$$
$$\sin(x+360) = \sin(x)$$



On a unit circle, a 360^0 angle corresponds to a circumference of 2*pi

Sine and Cosine

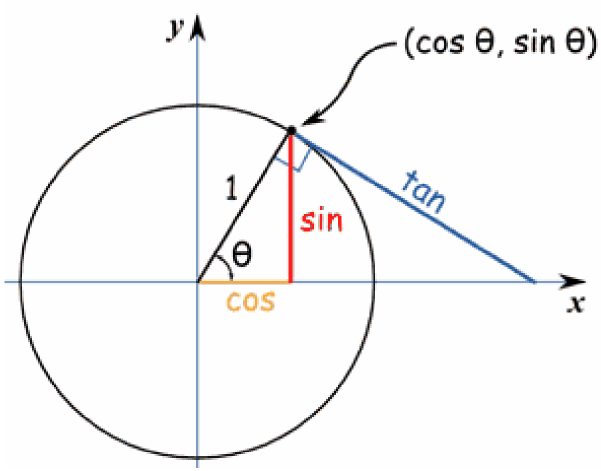
$$s(t) = a * sin(2\pi f * t + \theta)$$

a: amplitude

f: frequency

θ : phase

$$\cos(x) = \sin(x + \frac{\pi}{2}) \text{ or } \cos(x) = \sin(x + 90)$$



Inverse of sine and cosine: arcsine and arccosine

Given the sine/cosine values, they will yield the angle.

https://www.youtube.com/watch?v=z82I6u4DFTo

Polar Representation Of Periodic Signals

Euler's Formula

Complex numbers can capture the two axes of the coordinate system for the circle around which the vector rotates periodically – this is rather abstract but helps the notation enormously.

$$e^{-i\theta} = \boldsymbol{cos}(\theta) + \boldsymbol{i} * \boldsymbol{sin}(\theta) \quad \boldsymbol{i}=\sqrt{-1}$$

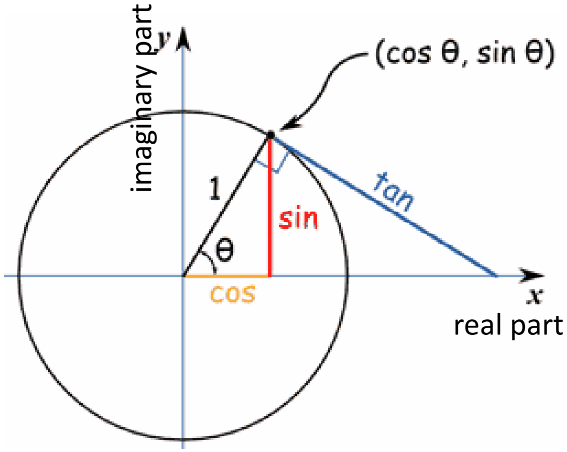
Therefore:

$$cos(\theta) = real(e^{-i\theta})$$
$$sin(\theta) = imag(e^{-i\theta})$$

An oscillation at a particular frequency can be described in a "polar representation":

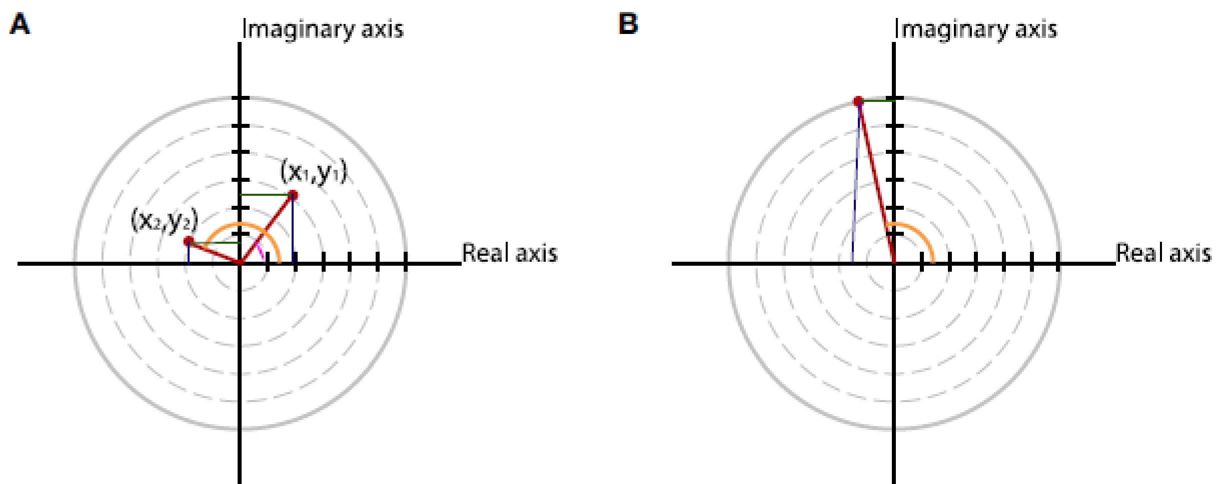
$$a * e^{-i2\pi f t}$$

a: amplitude
2π: circumference of unit circle
f: frequency
t: time



The Polar Representation Of Periodic Signals

Convenient To Compare Periodic Signals



signal 1 = $x_1 + iy_1 = A_1 e^{\phi_1}$
signal 2 = $x_2 + iy_2 = A_2 e^{\phi_2}$

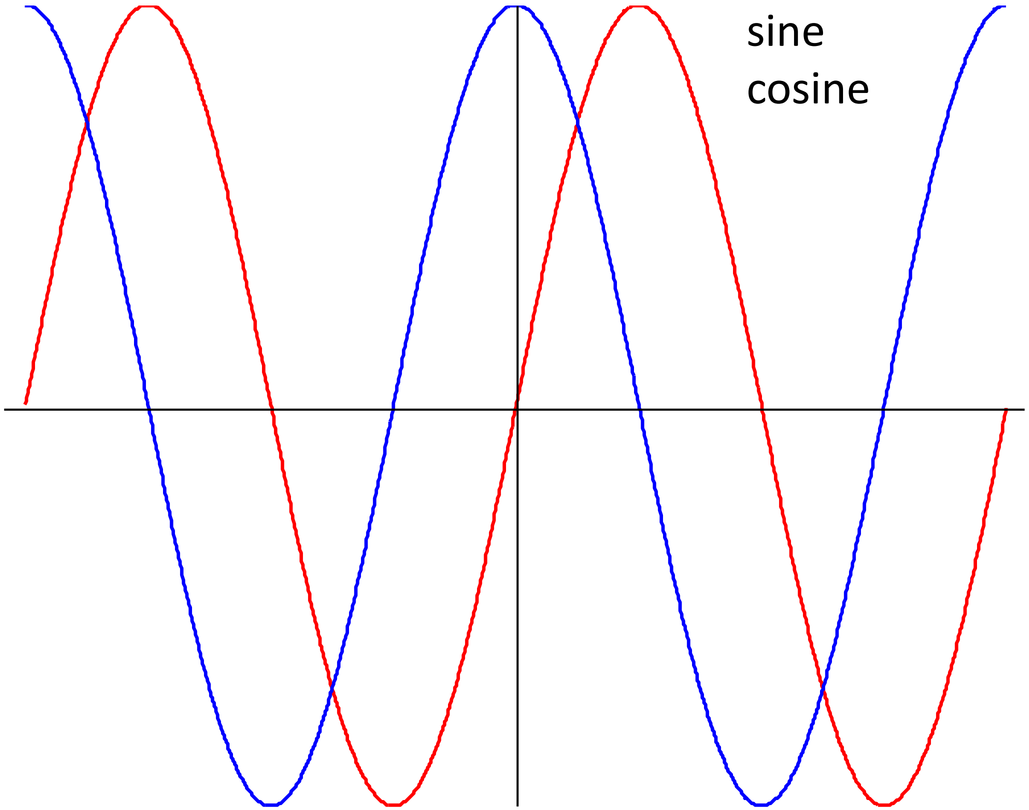
signal 1 * conj(signal 2) =
$(x_1 + iy_1) * (x_2 - iy_2) = (x_1x_2 + y_1y_2) + i(y_1x_2 - y_2x_1) =$
$A_1e^{\phi_1} * A_2e^{-\phi_2} = A_1A_2e^{(\phi_1 - \phi_2)}$

FIGURE 2 | Using polar coordinates and complex numbers to represent signals in the frequency domain. (A) The phase and amplitude of two signals. **(B)** The cross-spectrum between signal 1 and 2, which corresponds to multiplying the amplitudes of the two signals and subtracting their phases.

Bastos & Schoeffelen, Front Syst Nsc 2016
https://www.frontiersin.org/articles/10.3389/fnsys.2015.00175/full

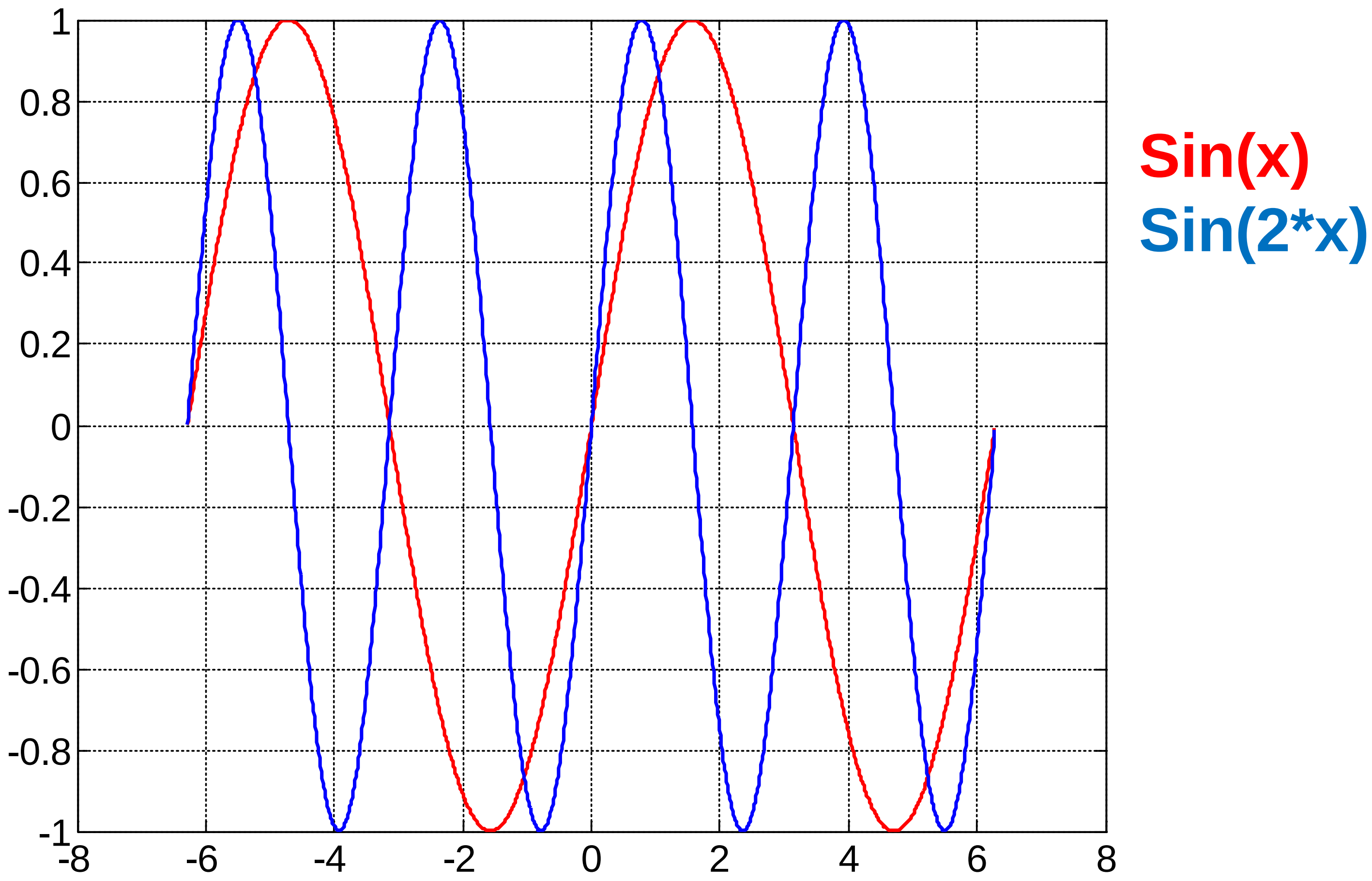
Sine and Cosine Are Orthogonal to Each Other

(at a given frequency)



$$\int \sin(f * x)\, cos(f * x) dx = 0$$

Sine/Cosine At Integer Frequency Intervals Are Orthogonal



$\int \sin(m * f * x)\, sin(n * f * x) dx = 0$ for integer m, n

Entering the Frequency Domain:
Fourier Transform in Words

What you want:

You've got a signal consisting of N sample points (equidistant).
You want to know which frequencies contribute to the signal, and how much.

In other words:
You want to describe your signal as a linear combination of sines and cosines, ideally of orthogonal basis functions made up of sines and cosines.

What you've got:

With N samples, you can estimate at most N independent parameters.

You cannot estimate frequencies above half of the sampling frequency SF (Nyquist).

For a given frequency, sine and cosine are orthogonal,
i.e. 2 basis functions per frequency.

Entering the Frequency Domain:
Fourier Transform in Words

Divide the frequency range 0 to SF/2 evenly into N/2 frequencies.

For every frequency, create a sine and a cosine.

Use these (orthogonal) sines and cosines as your basis functions.

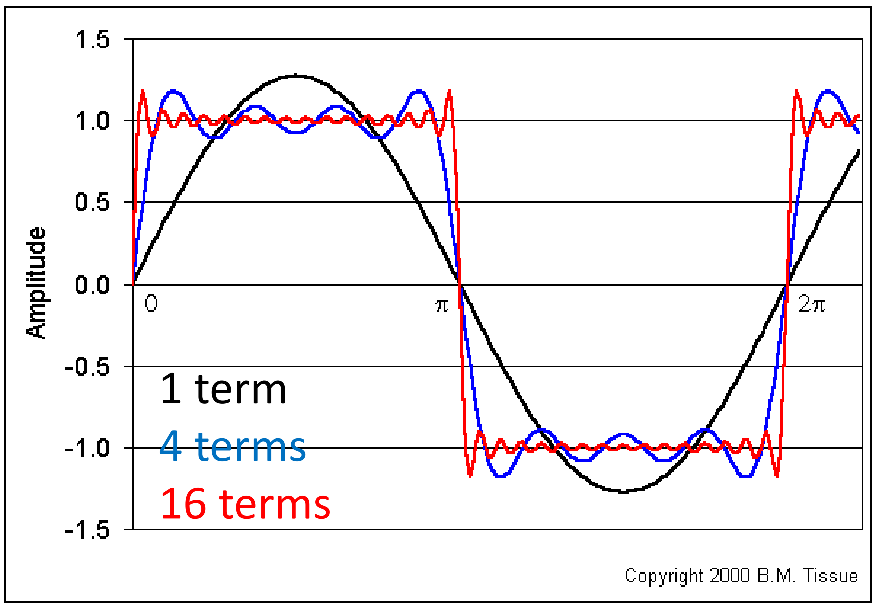
Project these basis functions onto your data, get the amplitudes for individual basis functions – that is your frequency spectrum.

Fast Fourier Transform (FFT): A fast algorithm to do this.

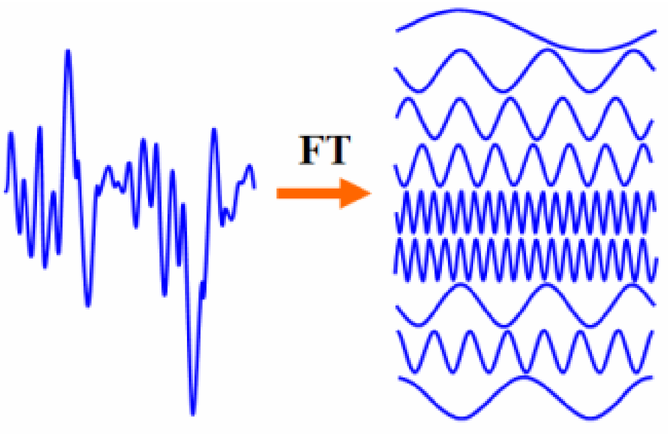
(I'm cheating a bit, assuming an appropriate N and ignoring the mean. But the principle is ok.)

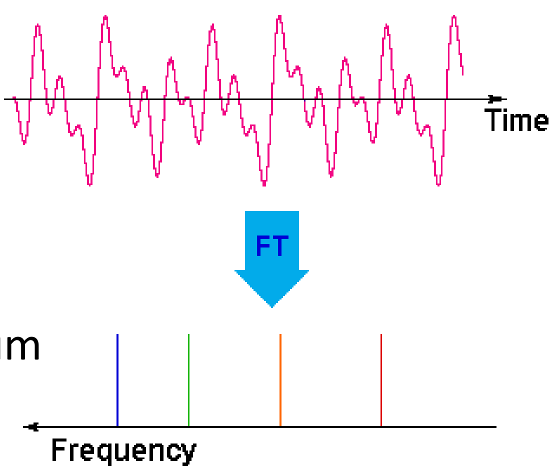
The Fourier (De-)Composition

Approximating a step function with Fourier terms



Decomposing signals into sine/cosine terms



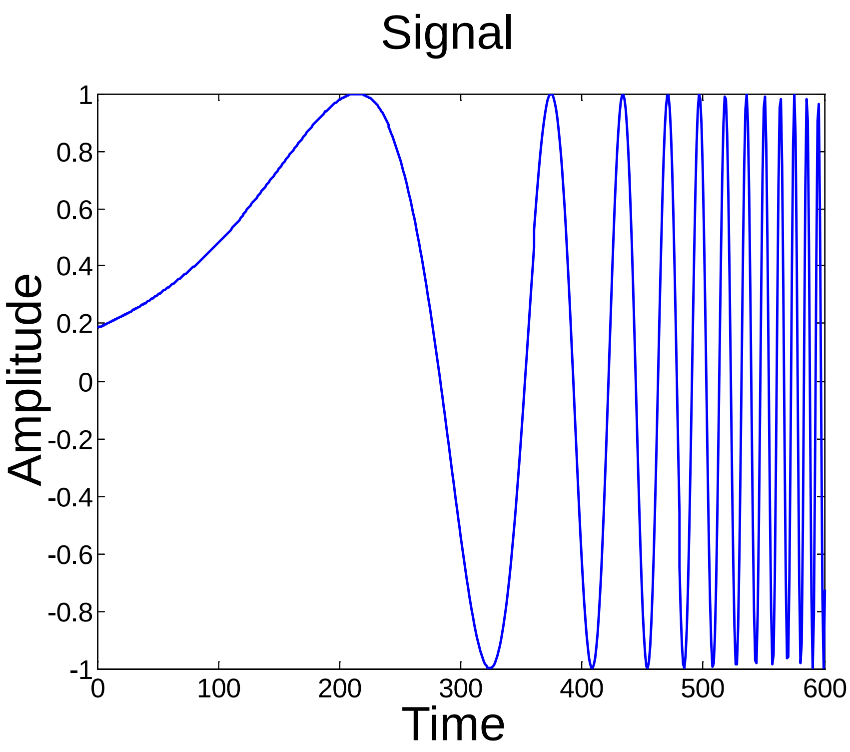


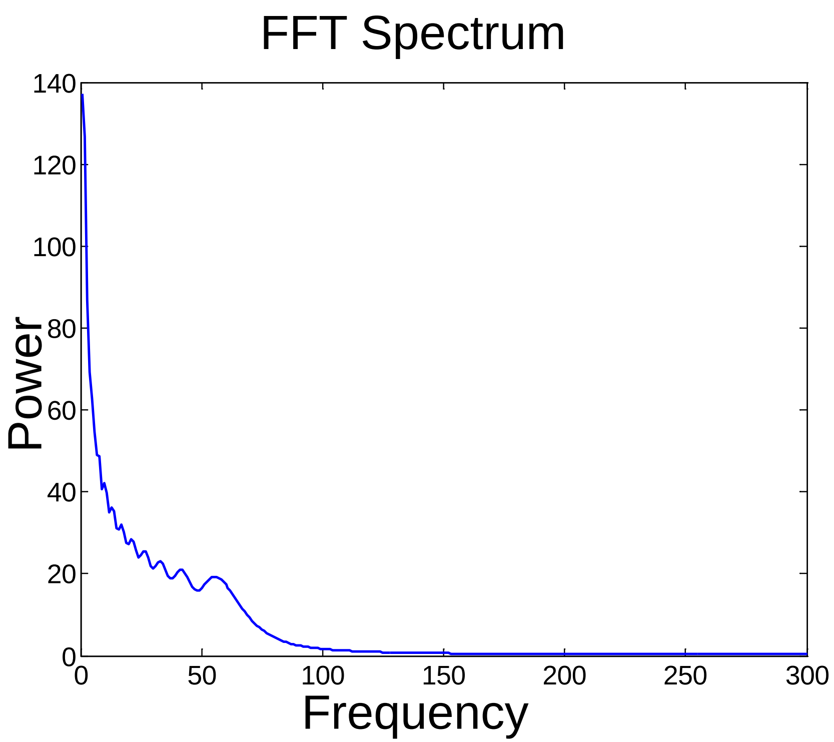
Frequency Spectrum

Motivation for Time-Frequency Analysis

Fourier Transform assumes sines and cosines with constant amplitudes across the whole time series ("stationarity").

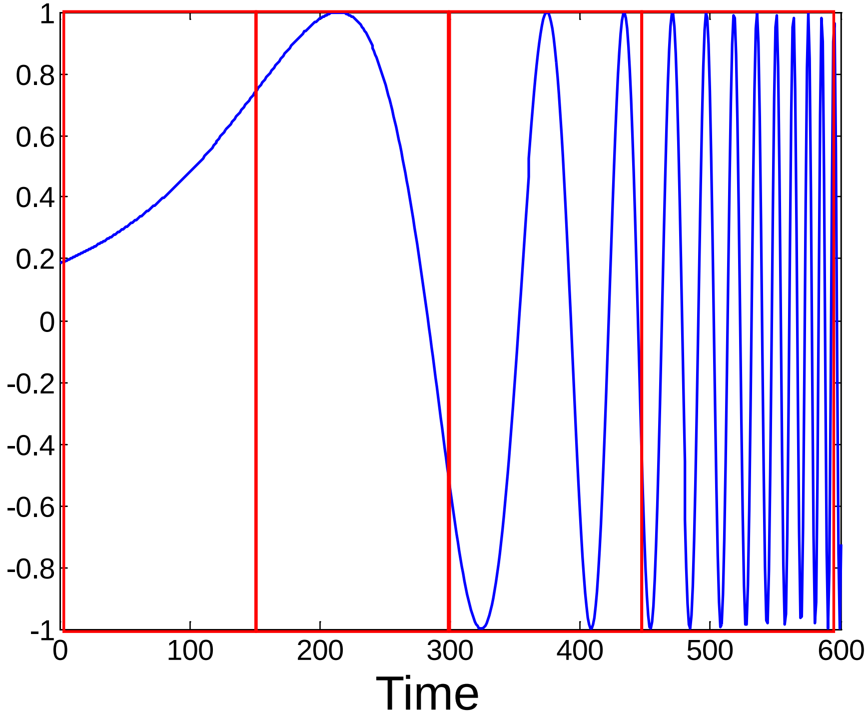
But what does an FFT mean for a signal like this?





Motivation for Time-Frequency Analysis

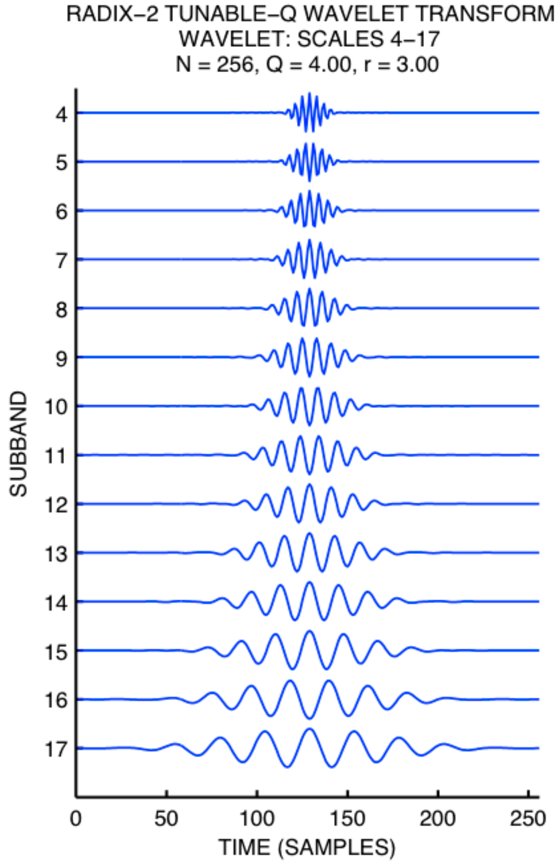
You could run separate FFTs for different (sliding) time windows:



But different window sizes are more or less optimal for different frequencies.
Run different FFTs with different window sizes for different frequency ranges? Ouff.

Time-Frequency Analysis: Wavelets

Wavelets provide an optimal trade-off between frequency and time resolution.

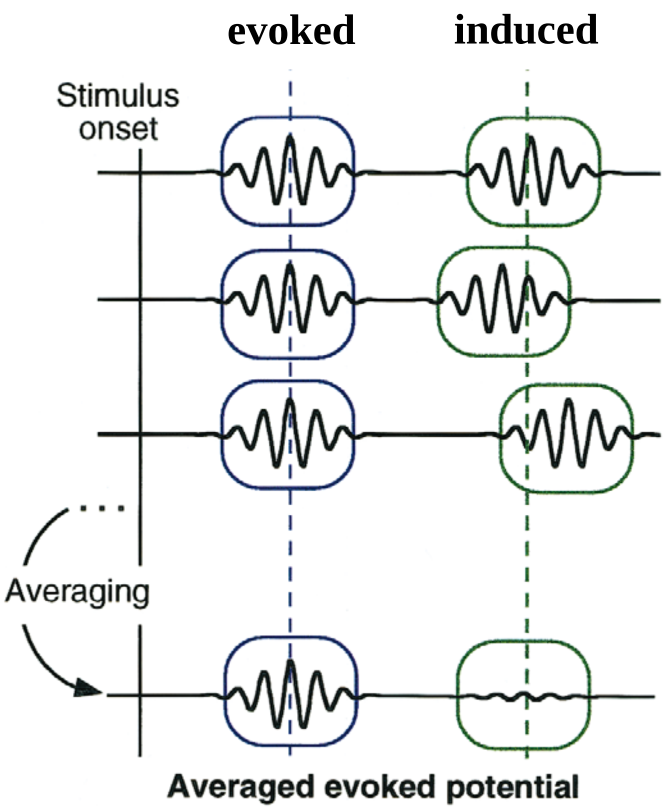


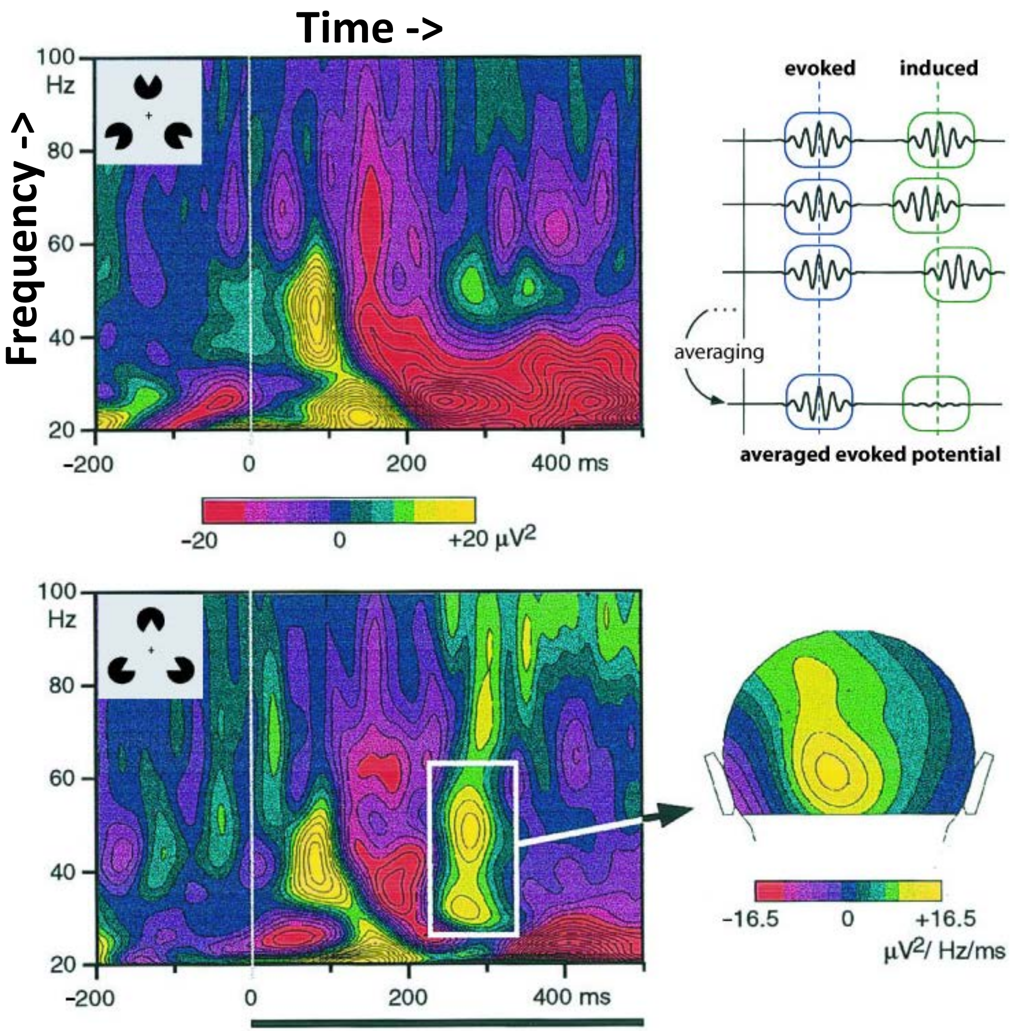
Time resolution decreases as frequency decreases

(wavelets are getting “broader”)

Wavelets are convolved with the data to give instantaneous amplitude and phase estimates for different frequency ranges.

Evoked and Induced Activity





Tallon-Baudry & Bertrand, TICS 1999

A Very Rough Rule of Thumb

One needs at least 2 cycles of a frequency to get a meaningful estimate (of amplitude, phase, etc.)

Duration (in ms) of 2 cycles at frequency f (in Hz): 2*1000/f

1 Hz: 2000 ms = 2 s

10 Hz: 200 ms = 1/5 s

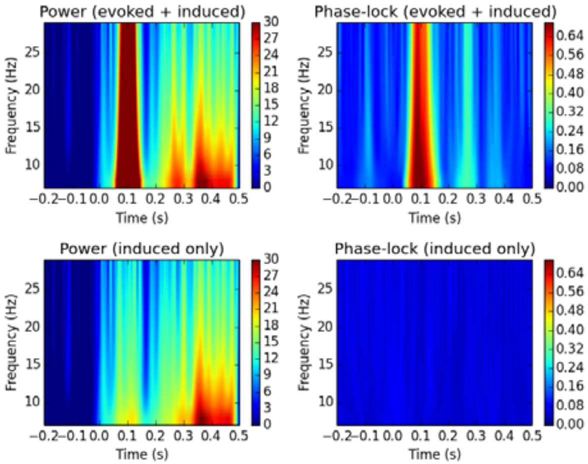
40 Hz: 50 ms = 1/20 s

100 Hz: 20 ms = 1/50 s

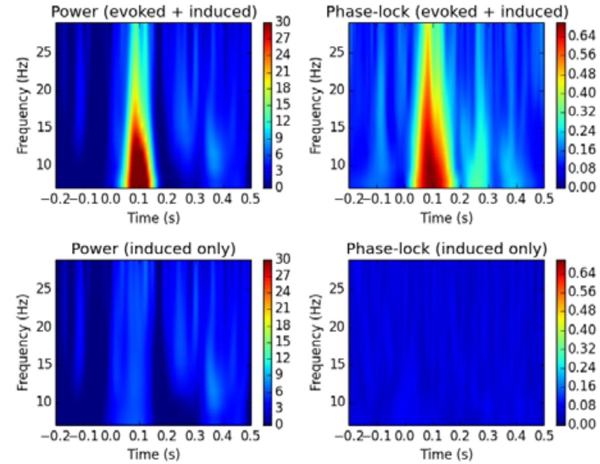
The lower the frequency, the longer the time window required to estimate the signal.

Effect of Number of Cycles

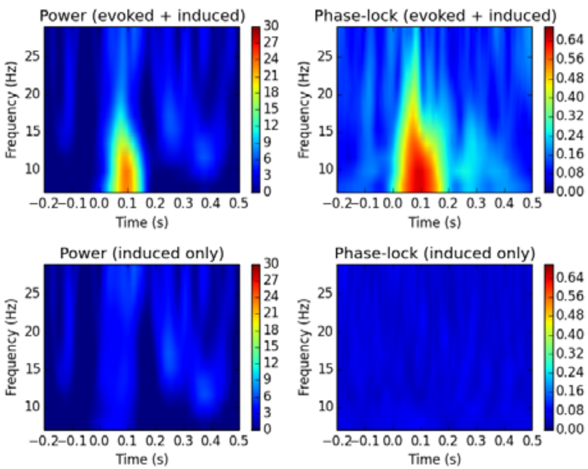
1 cycle



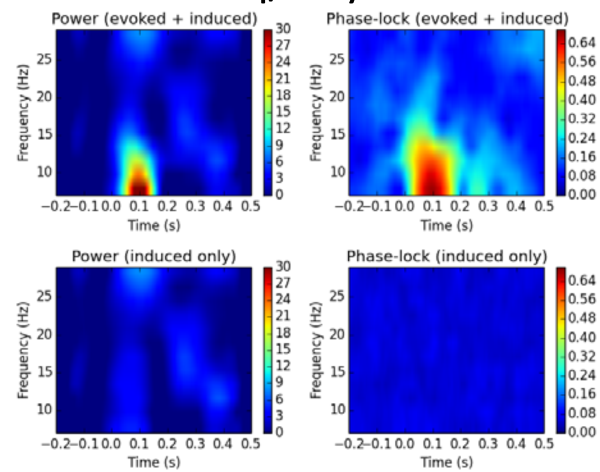
2 cycles



3 cycles



Freq/3 cycles



MRC

Cognition and Brain Sciences Unit

Single-Trial Analysis and Source Estimation

Computing the power of a signal is a non-linear transformation.

Linear transformations are associative:

$$T(a+b) = T(a)+T(b)$$

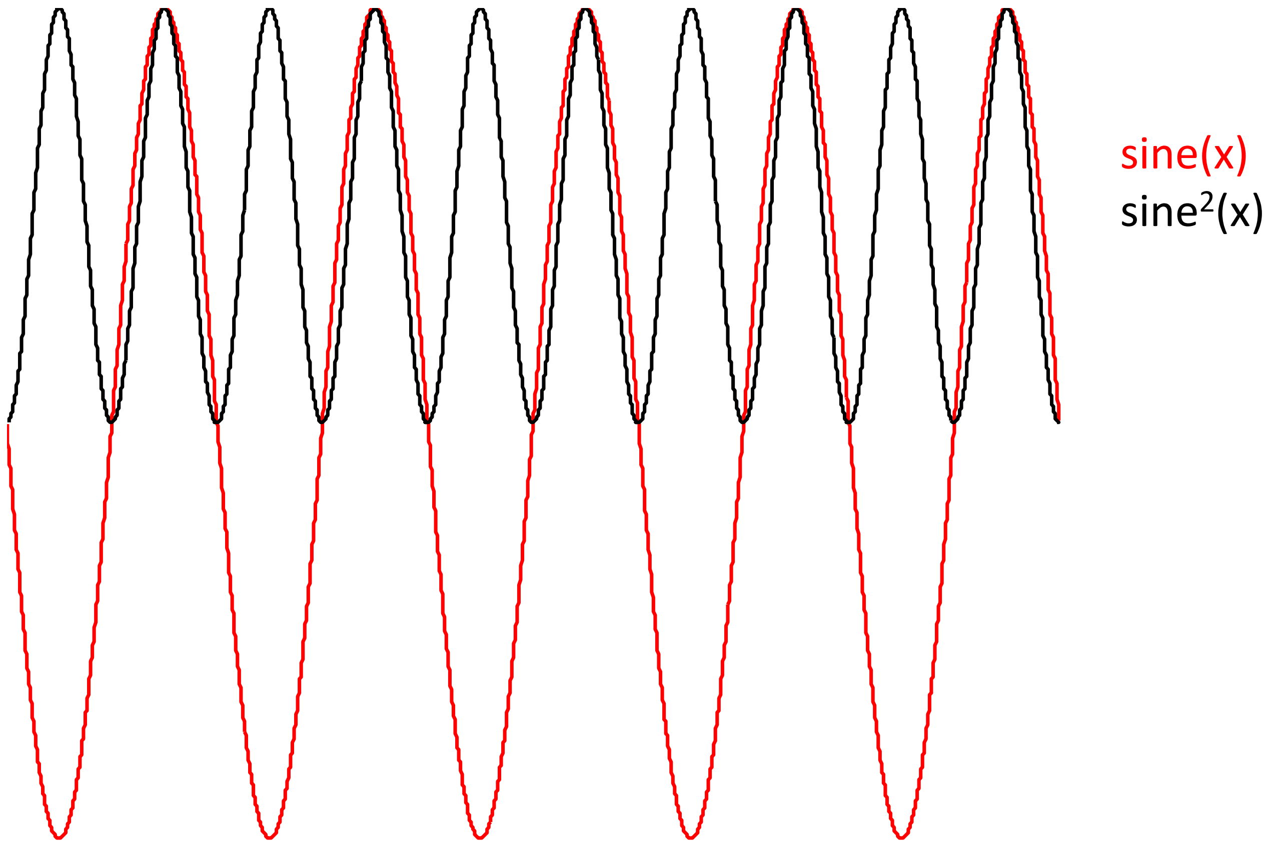
Therefore, the result is the same whether you apply a linear transformation before or after averaging your epochs.

Spectral power is non-linear!
If you want the average power, you have to compute power for individual epochs first, then average.

The noise level and a priori knowledge about sources will be very different for single trials compared to the average.

For example, a single/multiple dipole model may be justified for the average (e.g. auditory P1 etc.), but not for single trials.

Power Estimation Changes the Time Course



For example, the frequency spectrum for sine(x) and sine^2(x) are very different.

MRC
Cognition
and Brain
Sciences Unit

Brain Connectivity

Structural/Anatomical Connectivity:
Hardware links between brain regions
(e.g. DWI/DTI).

Functional Connectivity:
Statistical dependencies of activation between brain regions
(e.g. correlation, or spectral measures such as phase-locking and coherence).

Effective Connectivity:
Causal interactions of activation between brain regions
(Granger Causality, Dynamic Causal Modelling).

For example:
http://journal.frontiersin.org/article/10.3389/fnsys.2015.00175/full
http://www.sciencedirect.com/science/article/pii/S0165027012000817
http://www.ncbi.nlm.nih.gov/pubmed/21477655
http://online.liebertpub.com/doi/abs/10.1089/brain.2011.0008

Taxonomy Of Popular Functional Connectivity Methods

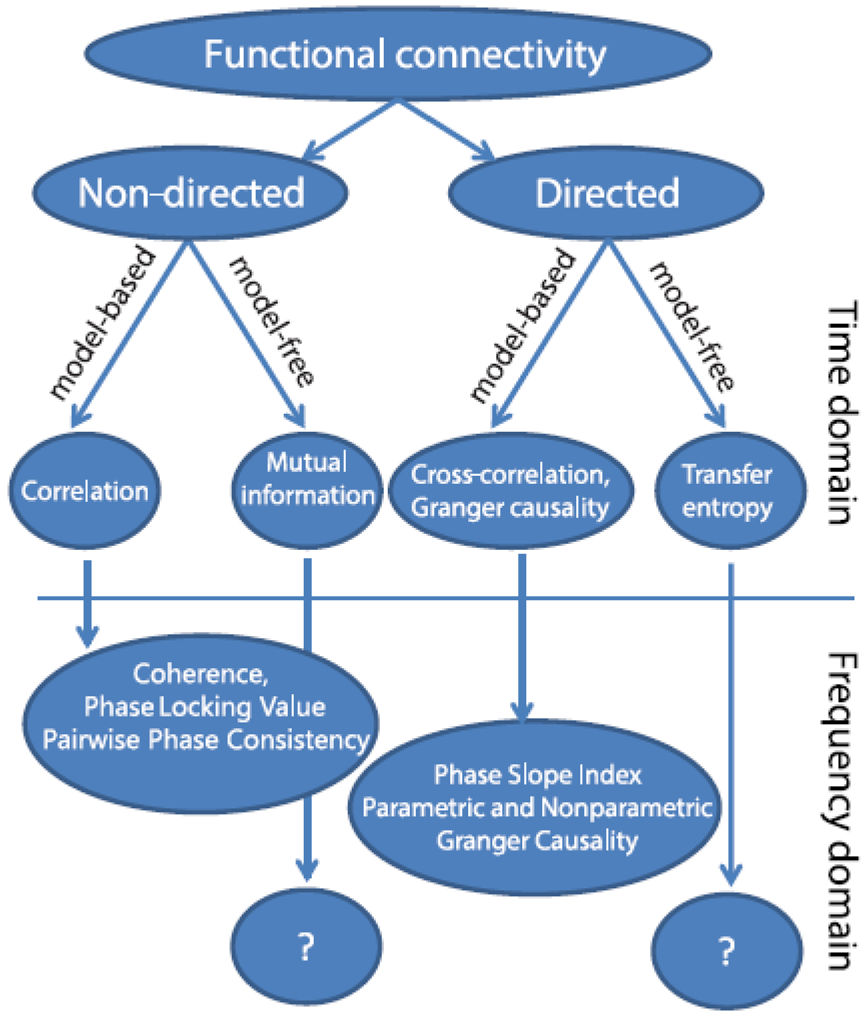


FIGURE 1 | A taxonomy of popular methods for quantifying functional connectivity.

Bastos & Schoeffelen, Front Syst Nsc 2016
https://www.frontiersin.org/articles/10.3389/fnsys.2015.00175/full

(Magnitude-Squared) Coherence

For two signals x(t) and y(t) at frequency f:

$$C_{xy}(\text{f}) = \frac{|G_{xy}(f)|^2}{G_{xx}(f)G_{yy}(f)}$$

$G_{xx}(f)$ *is power at f of x(t).*

$|G_{xy}(f)|^2$ *is cross – spectral density of* $x(t)$ and y(t).

$G_{xy}(f)$ is also called "Coherency" (and can be a complex number).

(MS-)Coherence yields the shared variance of two signals at a given frequency.

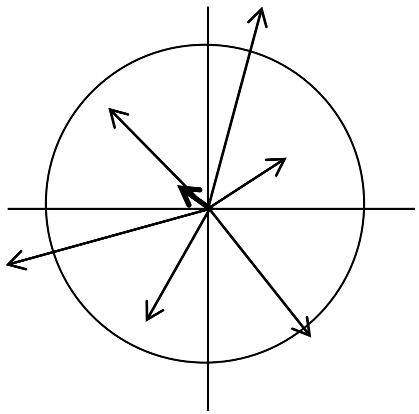
C_{xy}(f)=1: Signals perfectly coherent at frequency f.

C_{xy}(f)=0: Signals not coherent at all at frequency f.

This looks a bit like a correlation – but in this case it depends on amplitude and phase of the signals at frequency f.

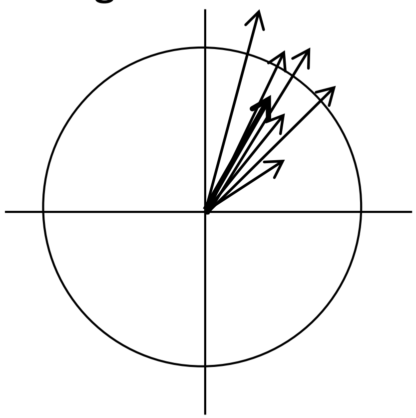
(Magnitude-Squared) Coherence

Low Coherence



Every vector represents the amplitude and phase of one signal
(e.g. phase difference between two regions across trials).

High Coherence



Coherence takes amplitude as well as phase consistency into account.
It can be interpreted as “amplitude-weighted phase-locking value”, i.e. trials with low amplitudes are given lower weight than those with higher amplitudes.

e.g., Bastos & Schoeffelen, Front Syst Nsc 2016
https://www.frontiersin.org/articles/10.3389/fnsys.2015.00175/full

Phase-Locking – Use Only Phase, Ignore Amplitude

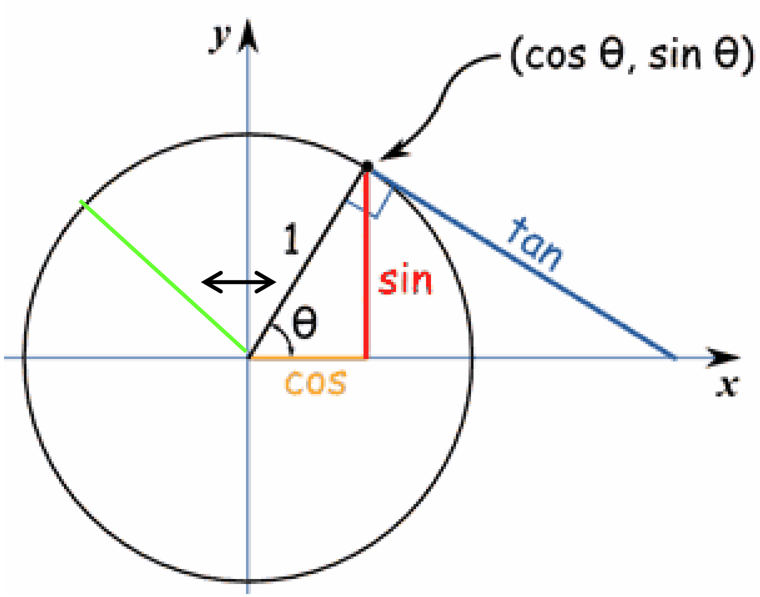
$$s(t) = a * sin(2\pi f t + \theta)$$

a: amplitude

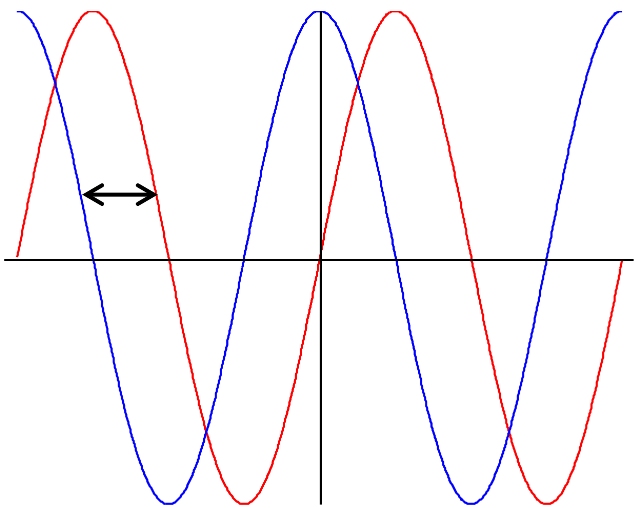
f: frequency

θ : phase

Phase difference in frequency domain



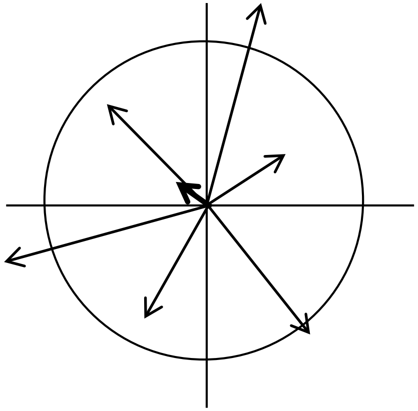
Phase difference in time domain



Lachaux et al., HBM 1999: http://www.ncbi.nlm.nih.gov/pubmed/10619414

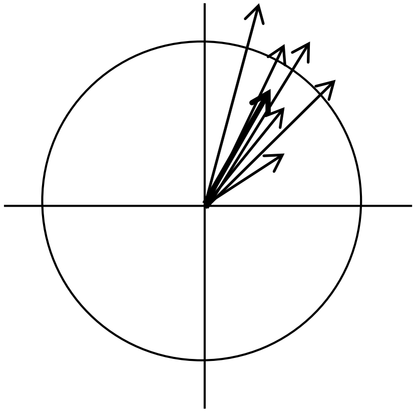
Phase-Locking vs Coherence

Low Coherence

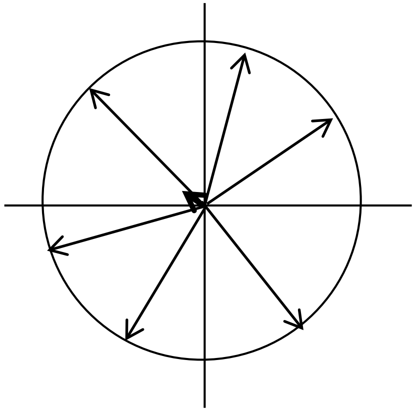


Every vector represents the amplitude and phase of one signal
(e.g. phase difference between two regions across trials).

High Coherence

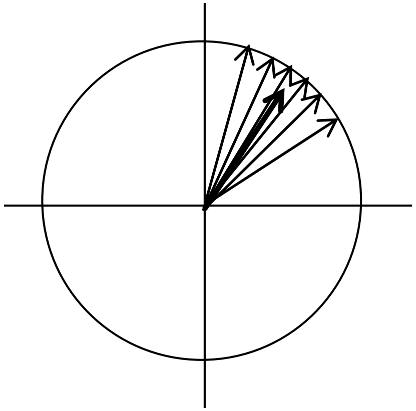


Low Phase-Locking



We are not interested in amplitude, and normalise all vectors to unit length.
The average vectors measure the phase-consistency across signals
(phase-locking value, PLV).

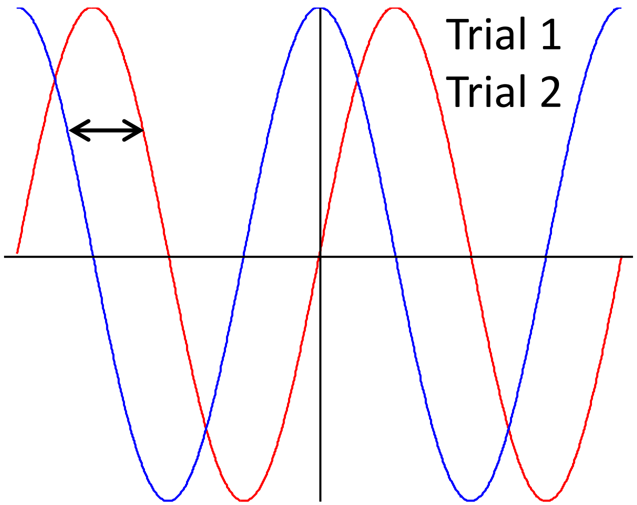
High Phase-Locking



Different Types of Phase-Locking

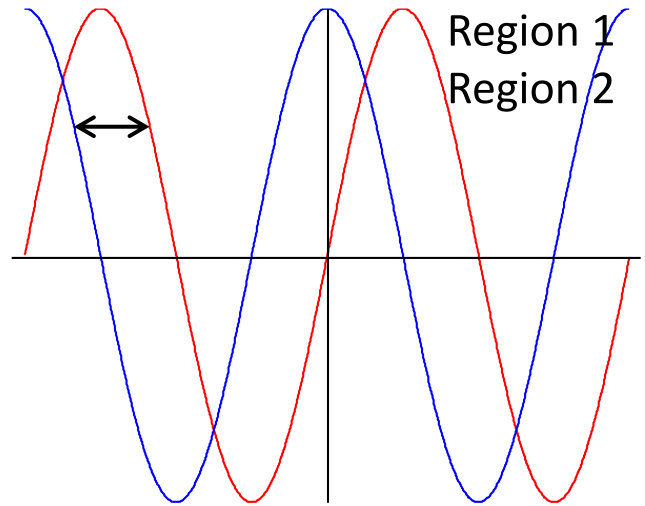
We ignore amplitudes, and are only interested in phase-relationships between two signal at a frequency f.

Inter-Trial Phase-Locking



Does the phase at a particular frequency remain stable across trials with one region?
(not connectivity)

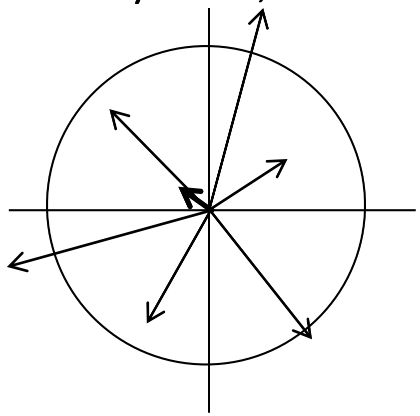
Inter-Regional Phase-Locking



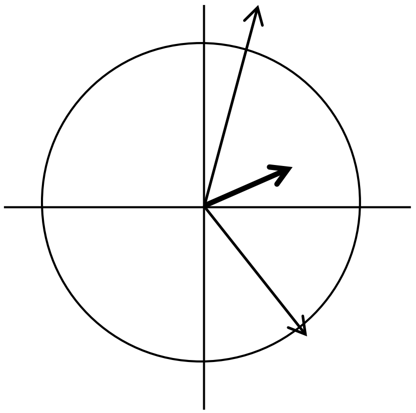
Does the phase difference between two regions at a particular frequency
remain stable across trials with one region?
(connectivity)

Sample Size / SNR Bias

Many trials, noise



Few trials, noise



Many connectivity metrics are positively biased (e.g. Coherence with values between 0 and 1), i.e. one gets positive values even in the presence of pure noise.
Importantly, the metric depends on the number of trials.

⇒ Plot metric for baseline data and different trials counts in your own data
⇒ Equalise trials counts between conditions
⇒ Baseline correction

This effects is relatively small for ~>50 trials:

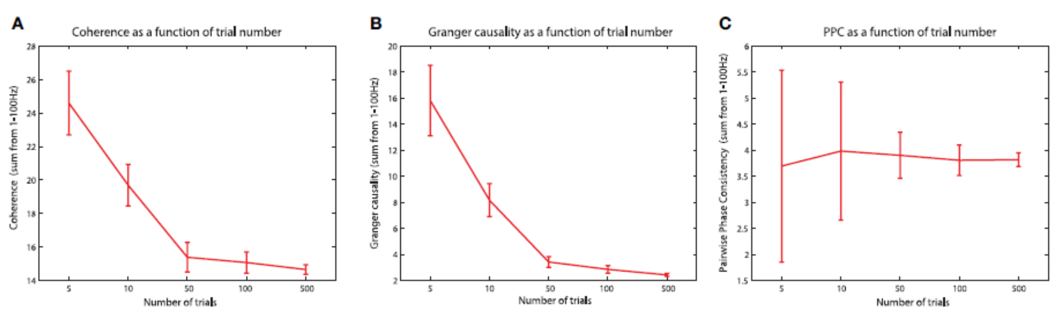
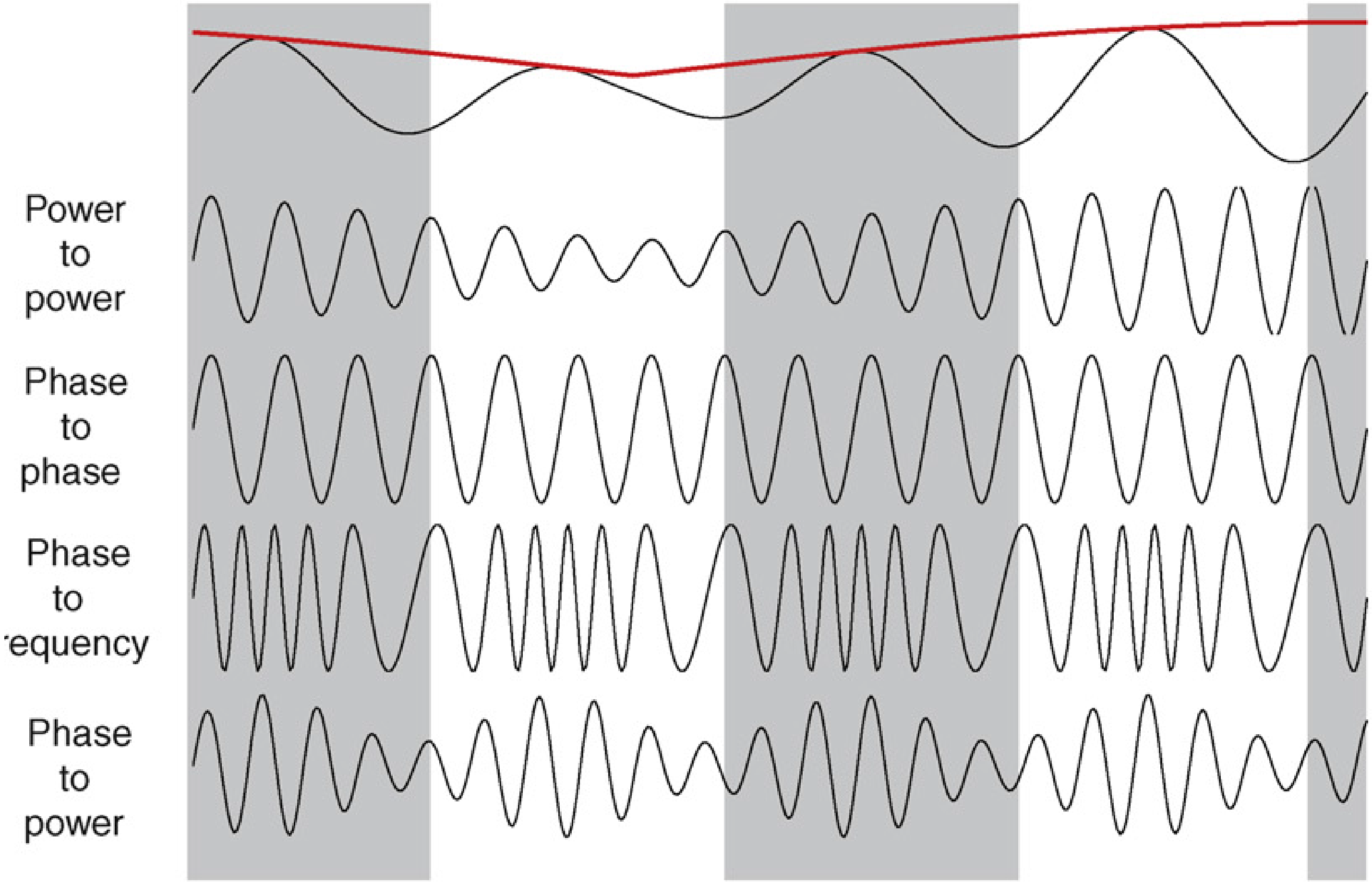


FIGURE 10 | Sample size bias for coherence and Granger causality estimates. (A–C) For each respective metric, simulations based on 5, 10, 50, 100, and 500 trials were run, and coherence **(A)**, Granger causality **(B)**, and PPC **(C)** were calculated. Each panel reflects the average ± 1 standard deviation across 100 realizations.

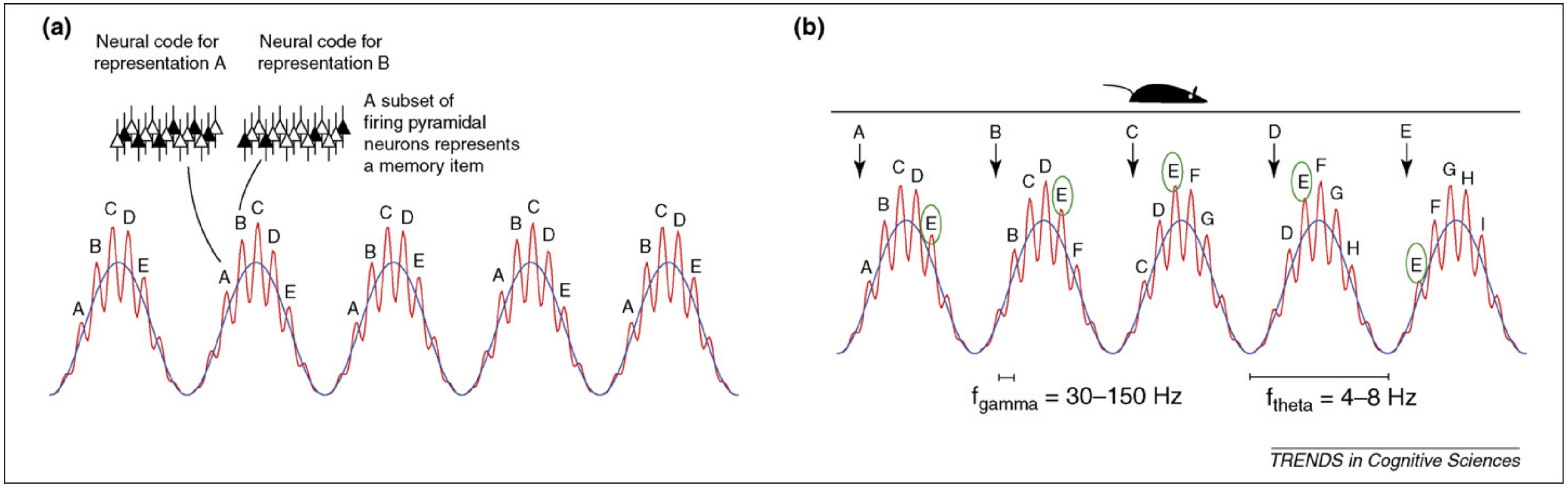
Bastos & Schoeffelen, Front Syst Nsc 2016
https://www.frontiersin.org/articles/10.3389/fnsys.2015.00175/full

Cross-Frequency Coupling



Jensen & Colgin, TICS 2007

For Example: Theta-Gamma Coupling

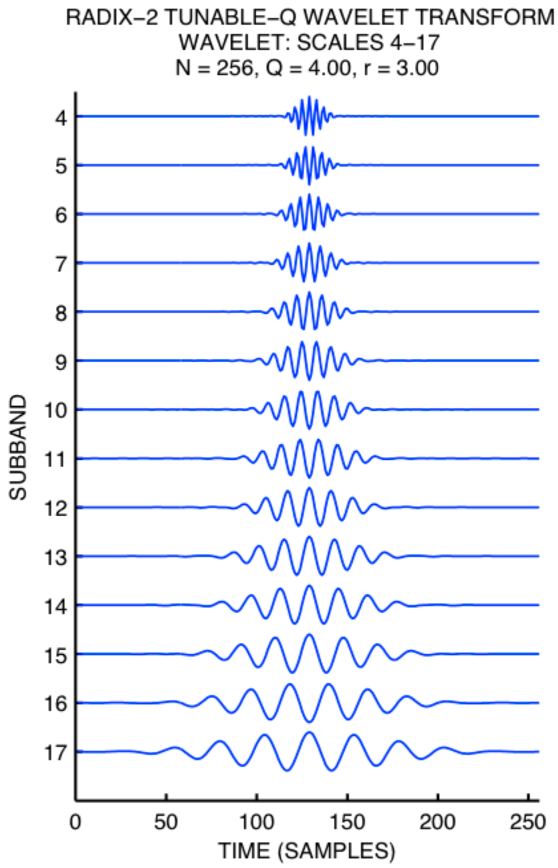


Jensen & Colgin, TICS 2007

Figure 2. Models proposing computational roles for cross-frequency interactions between theta and gamma oscillations by means of phase coding. (a) In a model for working memory, individual memory representations are activated repeatedly in every theta cycle [10] (reviewed in Ref. [11]). Each memory representation is represented by a subset of neurons in the network firing synchronously. Because different representations are activated in different gamma cycles, the gamma rhythm serves to keep the individual memories segmented in time. The number of gamma cycles per theta cycle determines the span of the working memory. (b) A model accounting for theta phase precession in rats. As a rat advances through an environment, positional information is passed to the hippocampus. This activates the respective place cell representations, which provokes the prospective recall of upcoming positions. In each theta cycle, time-compressed sequences are recalled: one representation per gamma cycle. Consider the firing of a cell participating in representation E. As the rat advances, this cell fires earlier in the theta cycle, thus accounting for phase precession. According to this scheme, the number of gamma cycles per theta cycle is related quantitatively to the phase precession [13].

Time-Resolved Connectivity

Spectral connectivity measures can be computed for separate time windows, or they can be computed continuously using wavelets or Hilbert transform (subject to general trade-off between frequency and time resolution)



Temporal resolution decreases as frequency decreases

(wavelets are getting “broader”)

Directed Functional Connectivity

Phase-Slope Index (PSI): For signals with a stable time delay, the phase in the frequency domain should depend linearly on frequency

Nolte et al, Phys Rev Let 2008, http://doc.ml.tu-berlin.de/causality/
Basti et al., NI 2018, https://www.sciencedirect.com/science/article/pii/S1053811918301897

Auto-regressive models, Granger Causality:

…in the time domain:

Predict the future of a signal based on the past of its own and other signals

…in the frequency domain:

- Partial Directed Coherence
- Directed Transfer Function

Bastos & Schoeffelen, Front Syst Nsc 2016, https://www.frontiersin.org/articles/10.3389/fnsys.2015.00175/full
Greenblatt et al., J Nsc Meth 2012, https://www.sciencedirect.com/science/article/pii/S0165027012000817
Haufe et al. NI 2013, https://www.sciencedirect.com/science/article/pii/S1053811912009469

And Beyond…

Most of the previously introduced measures are spectral measures, i.e. they are computed for specific frequencies (or frequency bands).

They rely on the assumption that brain signals can meaningfully be decomposed into "oscillations" or "frequency bands".

This is a big assumption, and may not be the case for all modalities, stimuli, tasks etc., or may not even be true in general.

Therefore…

Non-Spectral and Effective Connectivity

Granger Causality: Is one time series useful to predict another?
x(t) Granger-causes y(t) if past values of x(t) add information to past values of y(t) for predicting future values of y(t).

http://www.scholarpedia.org/article/Granger_causality
Multivariate Granger Toolbox: http://www.sussex.ac.uk/sackler/mvgc/
http://journal.frontiersin.org/article/10.3389/fnsys.2015.00175/full

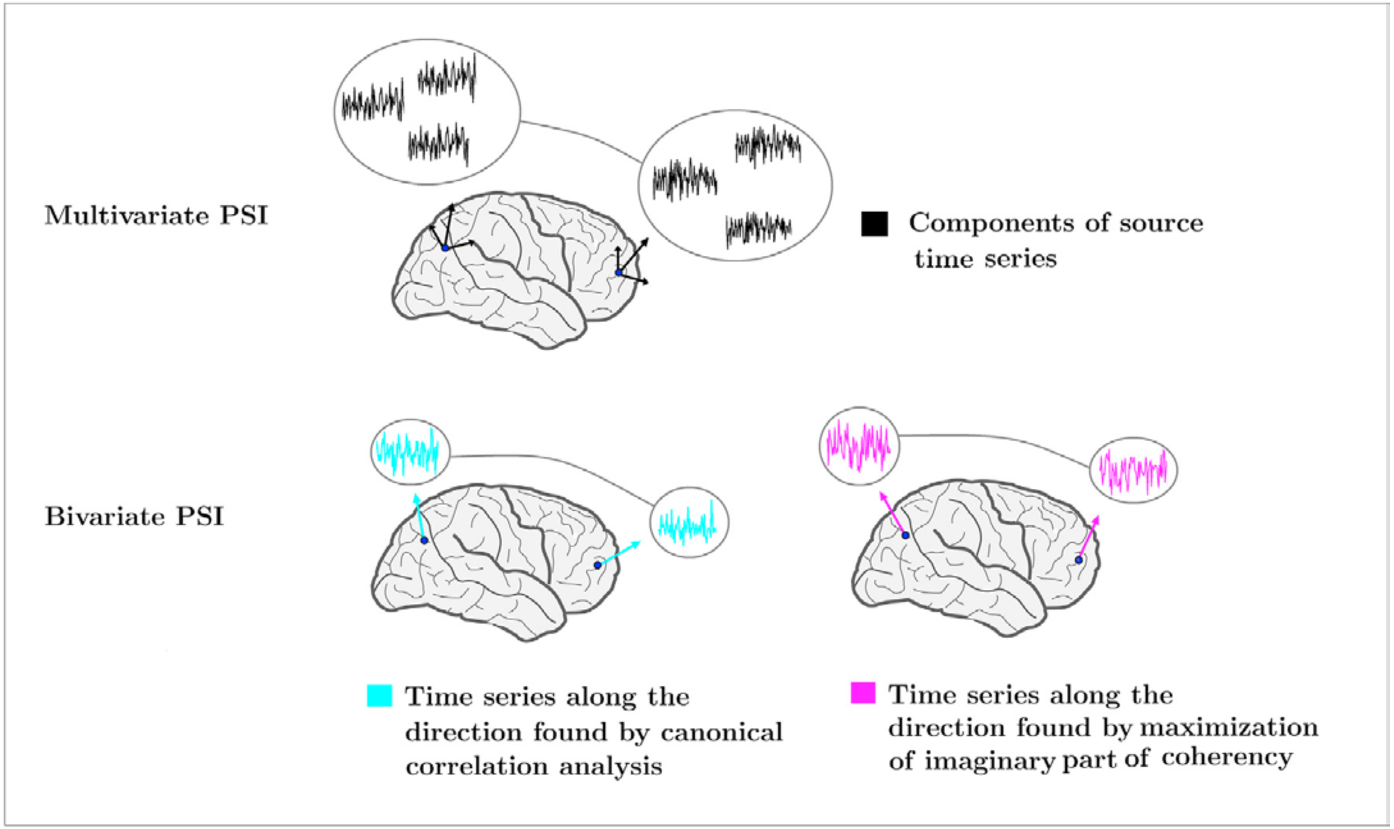
Structural Equation Modelling (SEM):
Models covariance structure of brain activation across brain regions (e.g. "path analysis").

Dynamic Causal Modelling (DCM):
Models brain dynamics across regions as differential equations, in combination with Bayesian parameter/model estimation.

http://www.scholarpedia.org/article/Dynamic_causal_modeling

Multi-Variate and Multi-Dimensional Connectivity

Currently, most connectivity methods use one time course per ROI. However, brain activity is multivariate, and there is potentially a lot of information lost by collapsing across vertices or voxels. “Multi-dimensional” methods are now emerging.



Basti et al., NI 2018, https://www.sciencedirect.com/science/article/pii/S1053811918301897

Also:

Basti/Nili et al., NI 2020, https://www.sciencedirect.com/science/article/pii/S1053811920306650

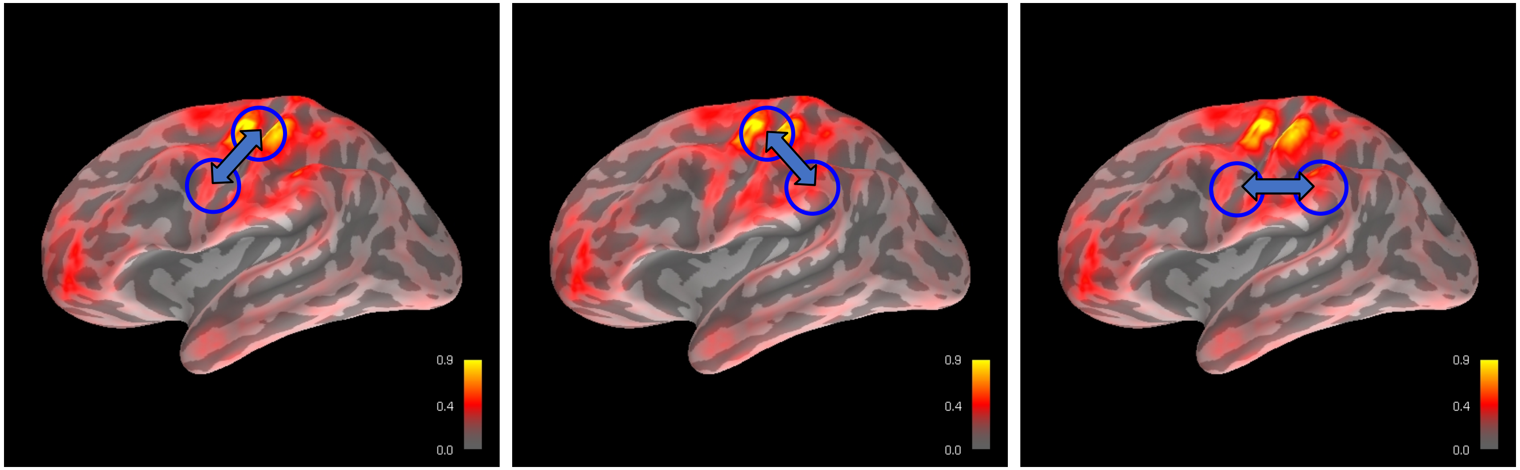
Anzellotti & Coutanche, T Cogn Sci 2018, https://pubmed.ncbi.nlm.nih.gov/29305206/

Basti et al., PLoS 2019, https://journals.plos.org/plosone/article/comments?id=10.1371/journal.pone.0223660

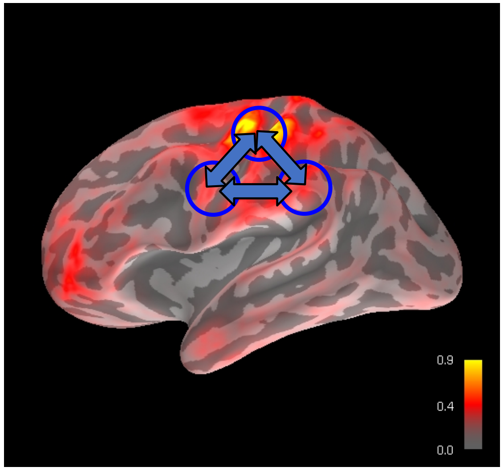
Spatial Resolution And Leakage Can Confound Connectivity Measures

Bivariate vs Multivariate Connectivity

Bivariate measures test one pair or regions at a time:

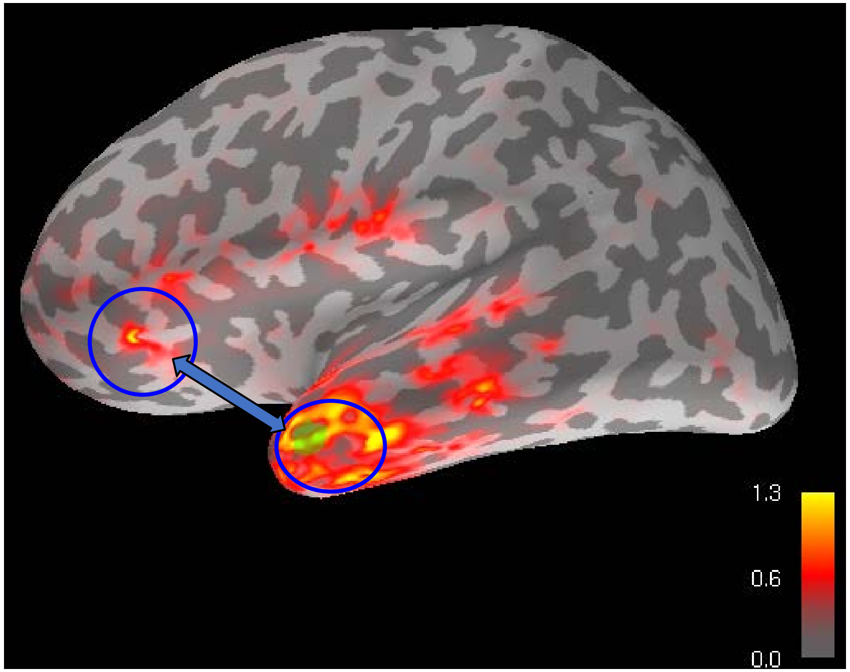


Multivariate measures test multiple regions simultaneously:

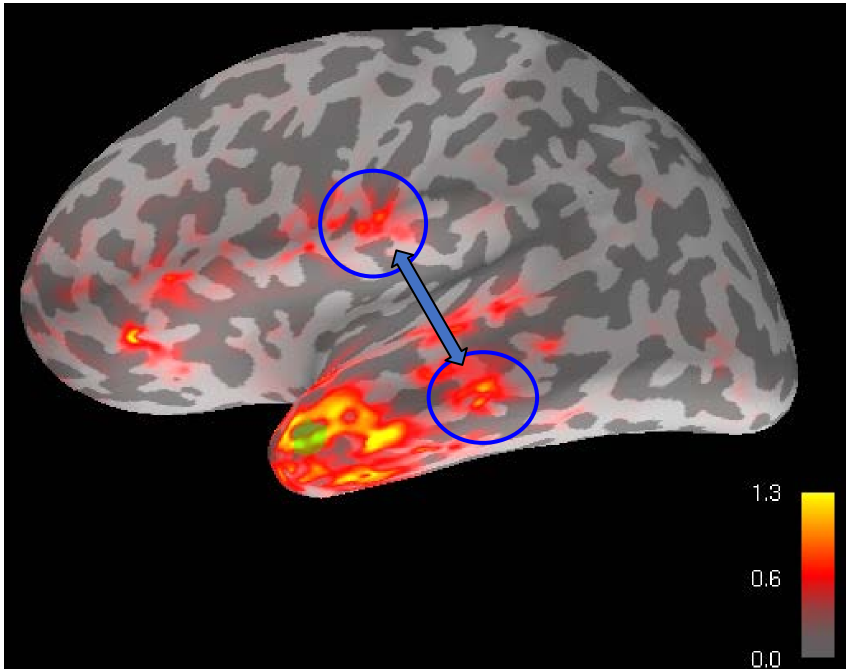


Field Spread / Point Spread

Connectivity between two regions may reflect cross-talk from one of the regions



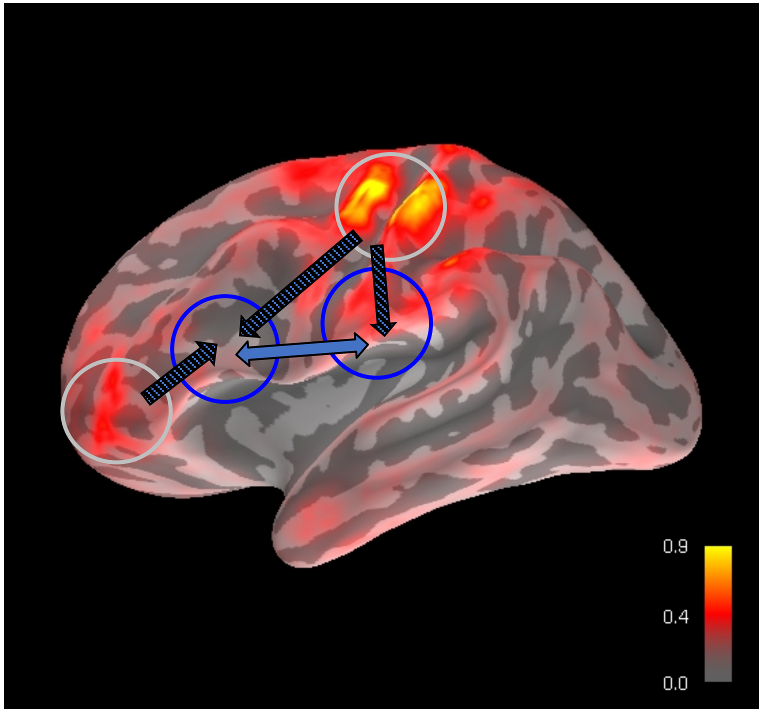
Connectivity between two regions may reflect cross-talk from a third region



Some connectivity measures can rule out "zero-lag" connectivity
(but they are then also insensitive to real zero-lag connectivity)

Field Spread / Point Spread

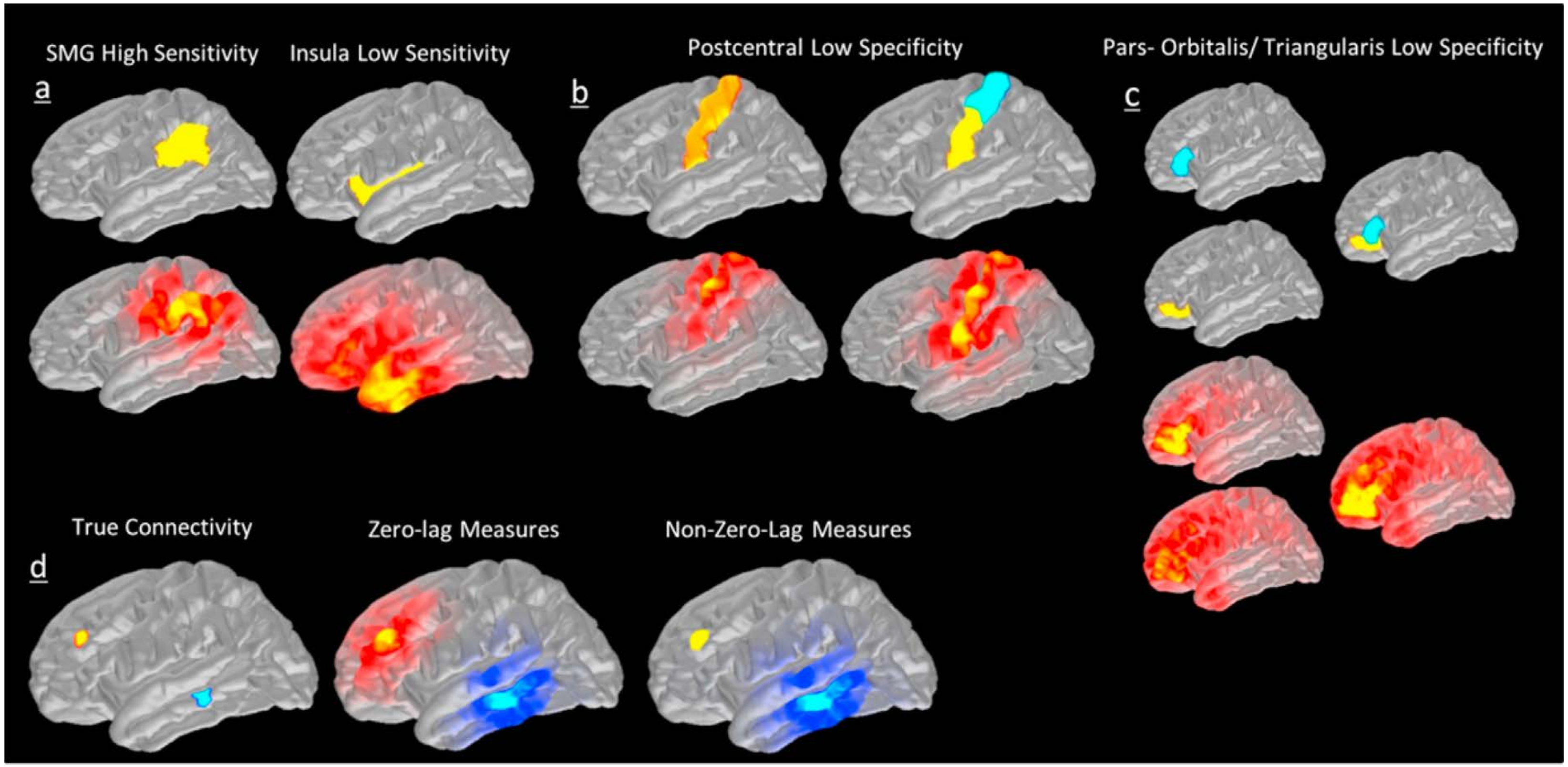
Connectivity between two regions may reflect cross-talk from several other regions



This is bad, and there is not much you can do –
except getting your model right in the first place, or use whole-brain analysis.

Leakage Can Produce Spurious Connectivity

(also at zero-lag)

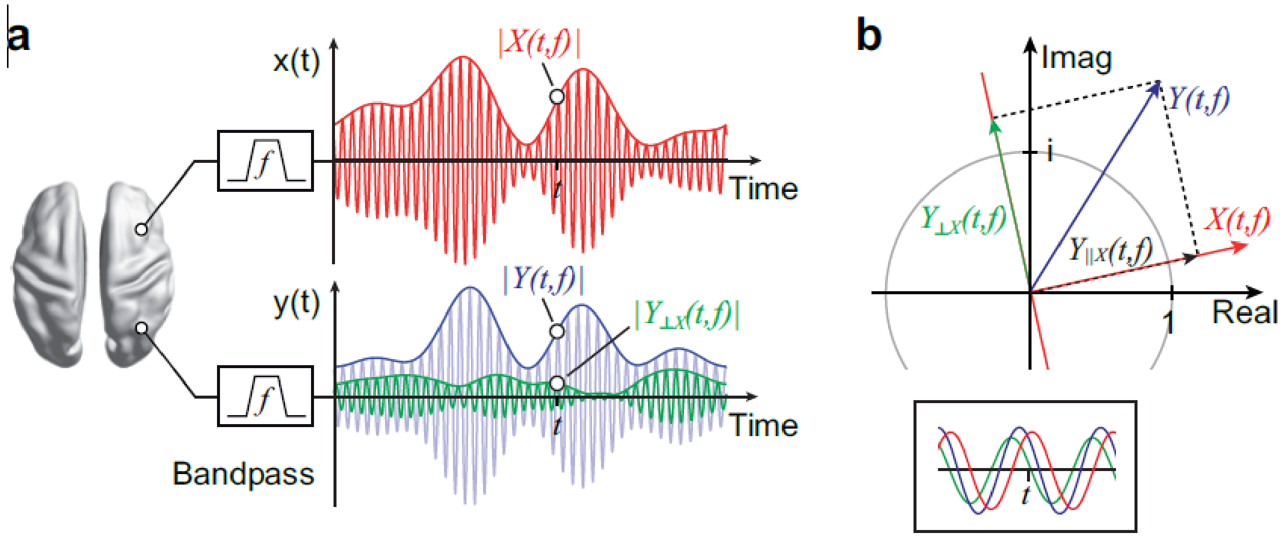


Farahibozorg, Henson, Hauk, NI 2018, https://pubmed.ncbi.nlm.nih.gov/28893608/
See also:
Palva et al., NI 2018, https://pubmed.ncbi.nlm.nih.gov/29477441/
Colclough et al. NI 2015, https://pubmed.ncbi.nlm.nih.gov/25862259/

One Possibility: Remove Zero-Lag Connectivity

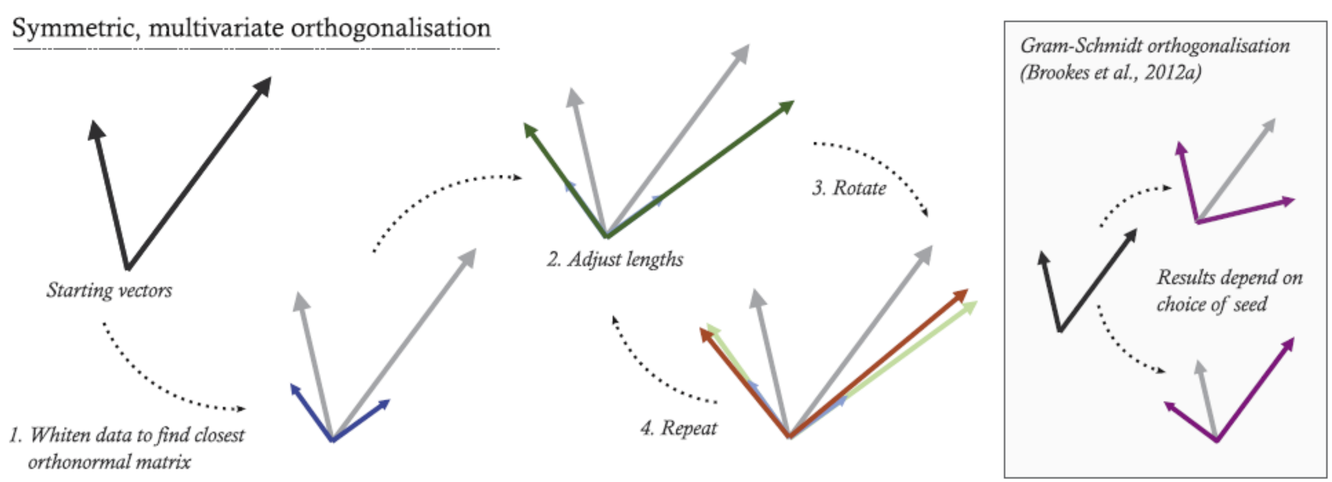
Orthogonalisation of time courses, Partial regression

Bivariate:



Hipp et al., Nat Nsc 2012, https://www.nature.com/articles/nn.3101

Multivariate:



Colclough et al., NI 2015, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4528074/

One Possibility: Remove "Zero-Lag" Connectivity

Imaginary Part of Coherency

In spectral connectivity measures like Coherence, only use the imaginary part of the signal, which is unaffected by zero-lag connectivity (phase differences of zero are only represented in the real part).

Ewald et al., NI 2012, https://pubmed.ncbi.nlm.nih.gov/22178298/

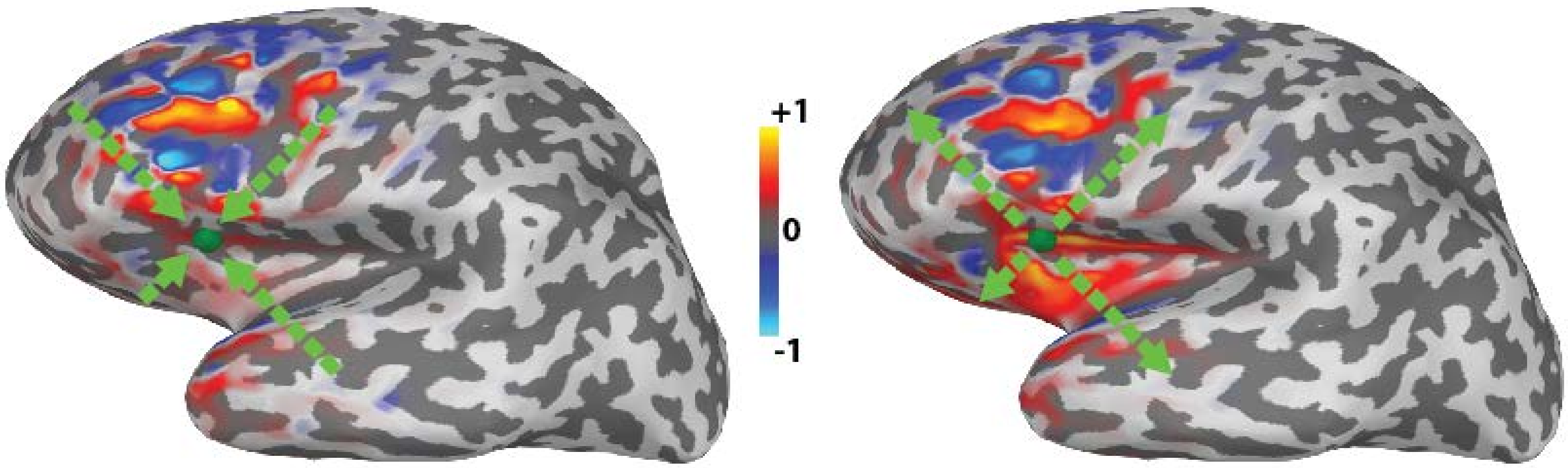
Pascqual-Marqui, arXiv 2007a and 2007b, https://arxiv.org/abs/0706.1776, https://arxiv.org/abs/0711.1455

"Non-zero-lag methods" may also ignore true zero-lag connectivity, e.g. for bilateral sources – one may through out the child with the bath water.

Spatial Resolution / Leakage:

Point-Spread and Cross-Talk

Cross-Talk Function (CTF) | Point-Spread Function (PSF)



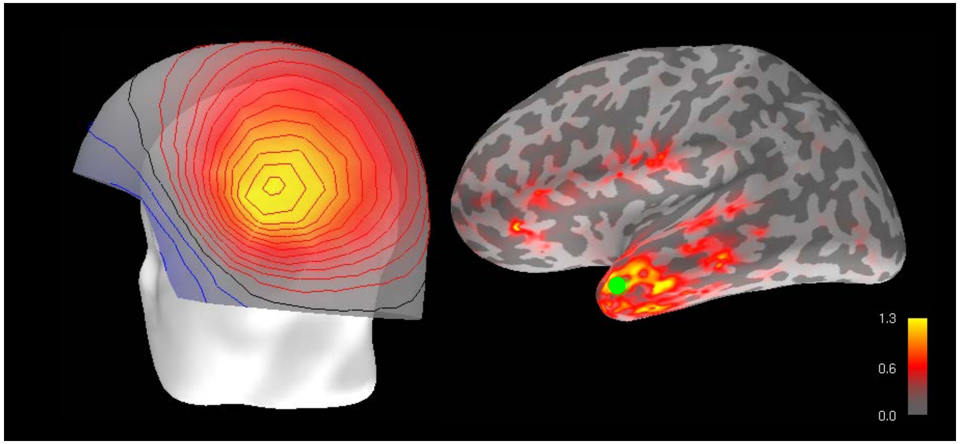
How other sources may affect the estimate for this source

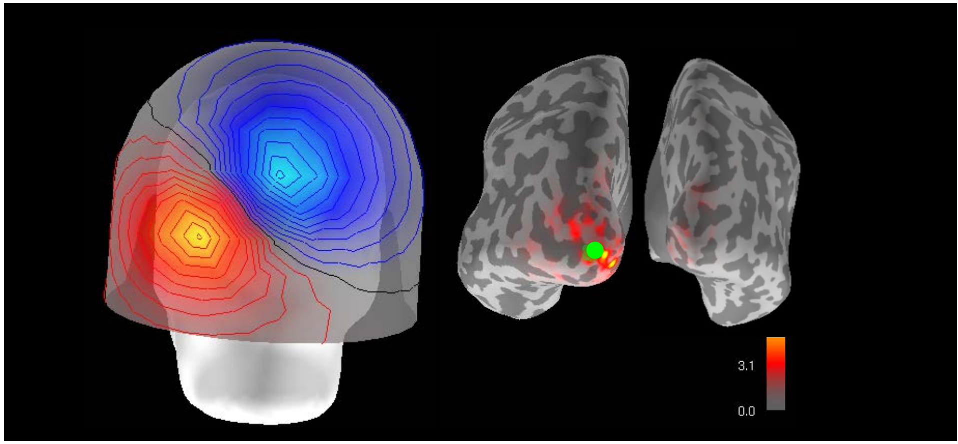
How this source affects estimates for other sources

e.g. Hauk/Stenroos/Treder, bioRxiv 2019
https://www.biorxiv.org/content/10.1101/672956v1

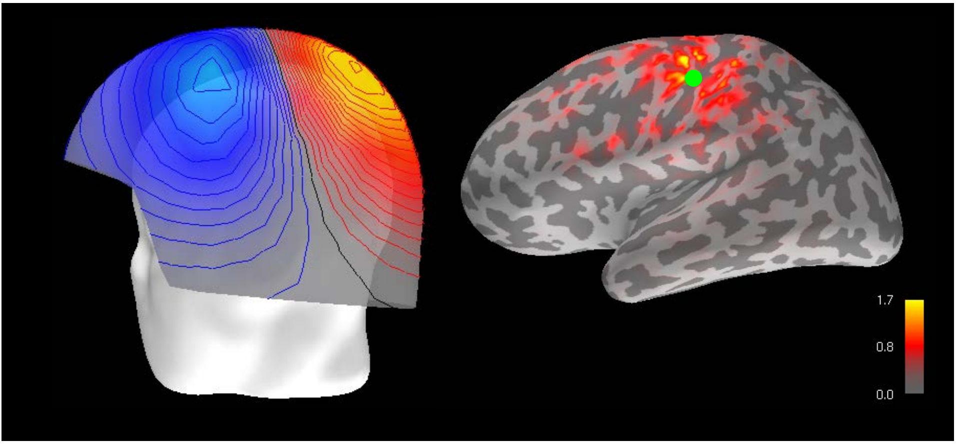
PSFs and CTFs for Some ROIs

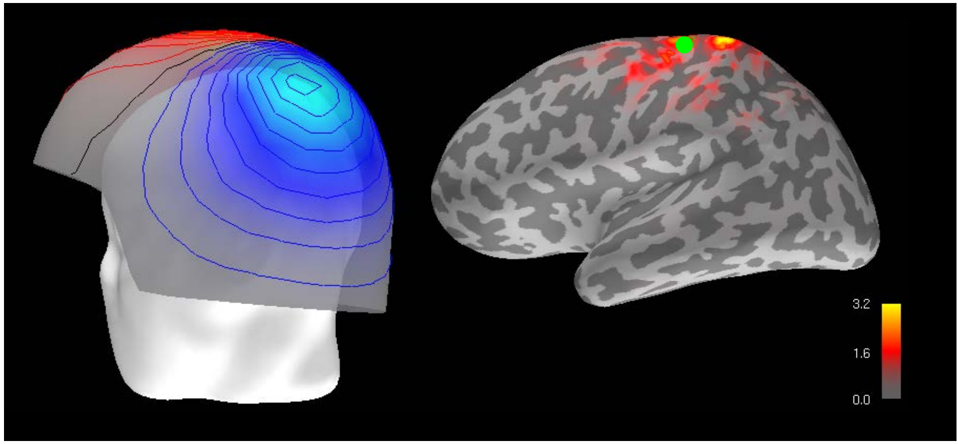
For MNE, PSFs and CTFs turn out to be the same





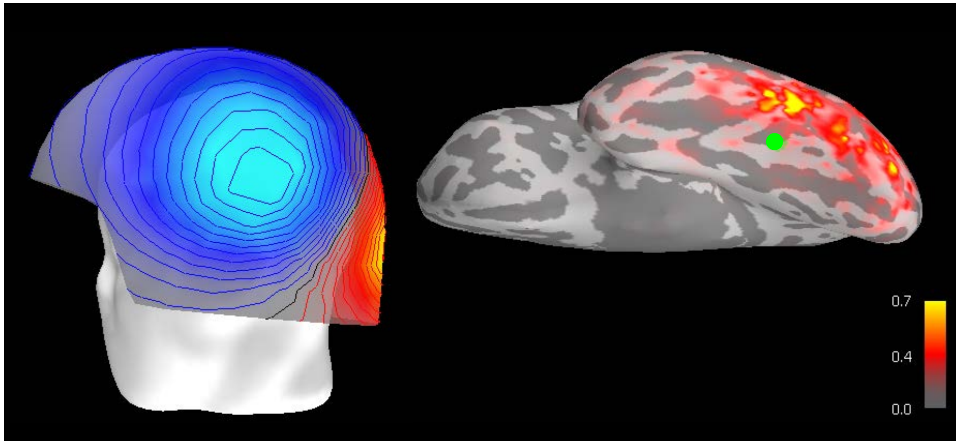
Good

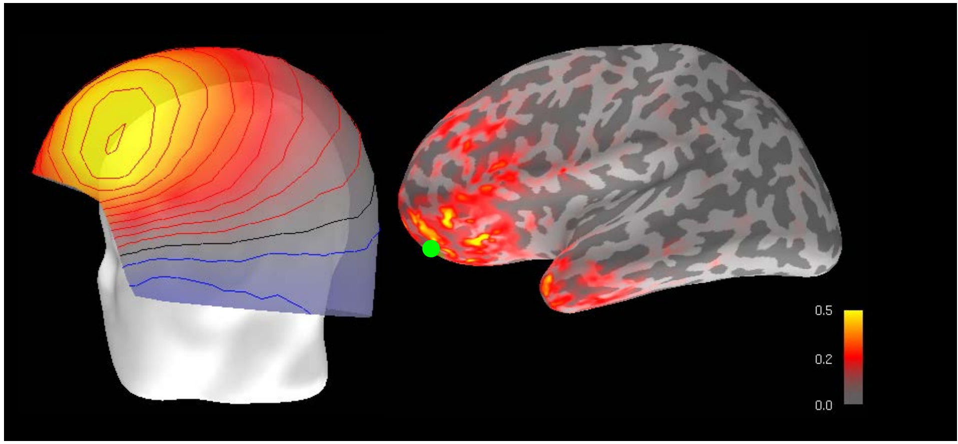




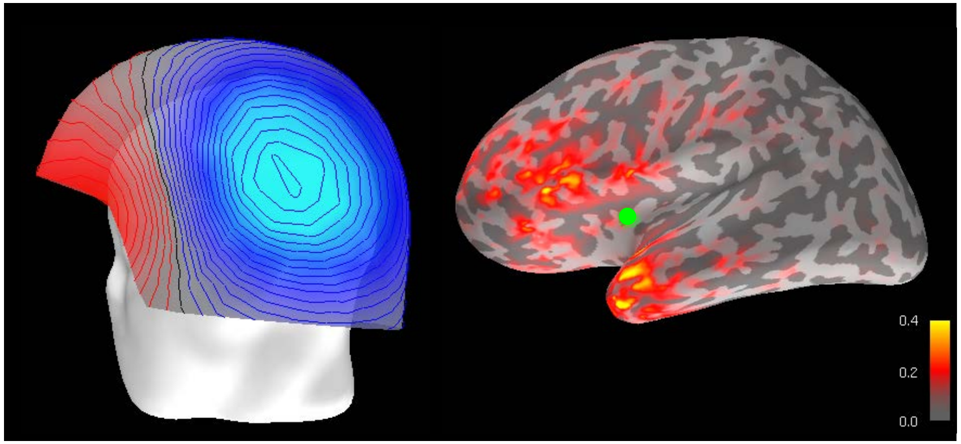
PSFs and CTFs for Some ROIs

For MNE, PSFs and CTFs turn out to be the same



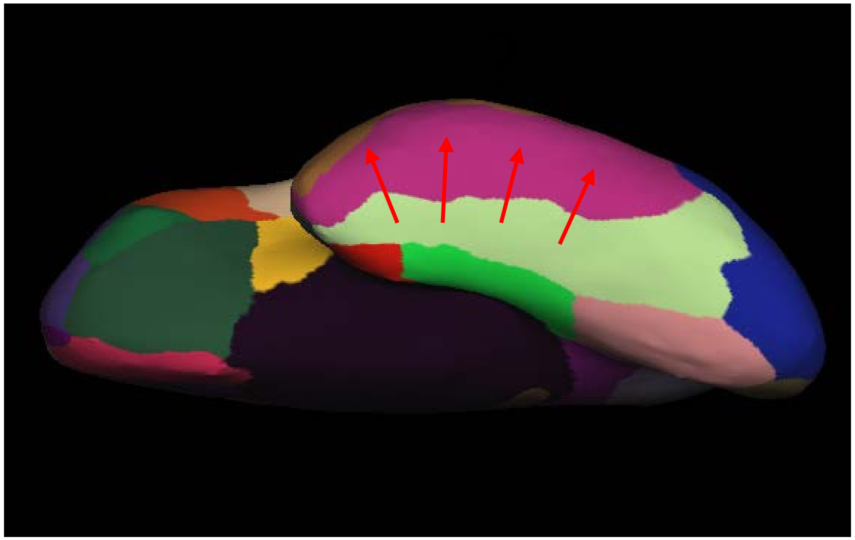


Less good

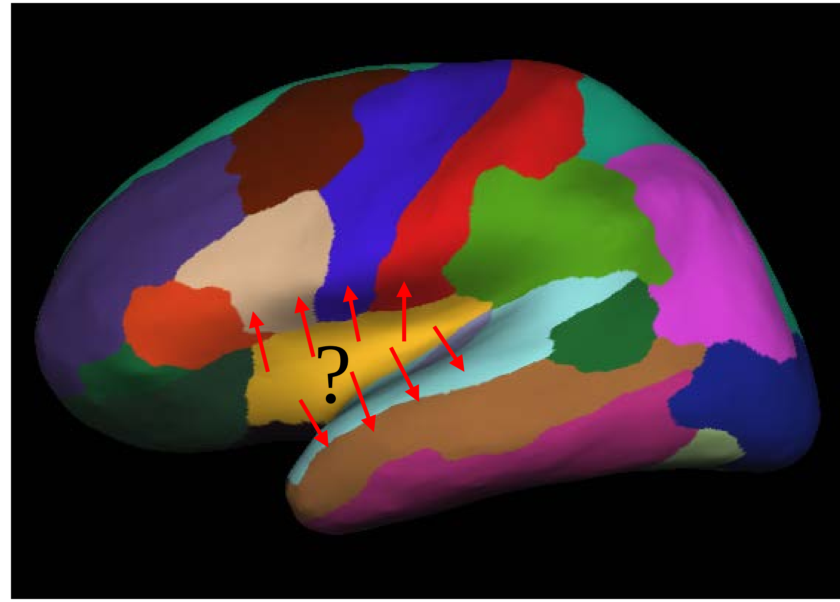


Localisation Bias Has Consequences for ROI analysis

PSFs/CTFs Can Tell You How It Looks Like



Desikan-Killiany Atlas parcellation



Adaptive cortical parcellation based on resolution matrix are possible:
Farahibozorg/Henson/Hauk NI 2018
https://pubmed.ncbi.nlm.nih.gov/28893608/