1. ERP: average frequency-band-specific power at each time point across trials
2. Power: amount of energy in frequency band. Amplitude2
   1. Interpretation: gamma (local processing), delta + theta (coordination of larger-scale network), alpha (correlate negatively w cortical activation – involved in inhibition), beta in motor regions (motor response), theta in prefrontal region (working memory, top-down cognitive control)
3. cross-frequency coupling: interaction btw oscillations at diff frequency bands. phase-phase coupling, amplitude-amplitude coupling, phase-amplitude coupling
   1. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3359652/
   2. phase-amplitude: low frequency phase reflects local neuronal excitability; high frequency power increases reflect
4. topological localisation (identify electrodes that show maximum effect) vs brain localisation (identify locations in brain that generated activity measured from scalp)
5. surface Laplacian technique: a type of spatial filter (to minimise overlap btw occipital + parietal responses)
   1. <https://www.sciencedirect.com/science/article/pii/S0167876015001749>
6. Distributed source imaging
   1. <ftp://www.besa.de/be/besa.de/demonstrations_and_tutorials/BESA%20Reserach%20Tutorial%204%20-%20Distributed%20Source%20Imaging.pdf>
7. Interpolation algorithms use weighted distance metric eg. nearest-neighbour, linear, spline
8. Artifacts that contaminate EEG data: blinks, muscle movements, brief amplifier saturations, line noise
   1. Removing blinks: independent components analysis, regression-based techniques
9. Microstate
10. Nyquist frequency – 1.5 of sampling rate

Time-frequency decomposition methods – set data point by weighted sum of surrounding data points, weights constructed to maximise selectivity to specific frequency bands

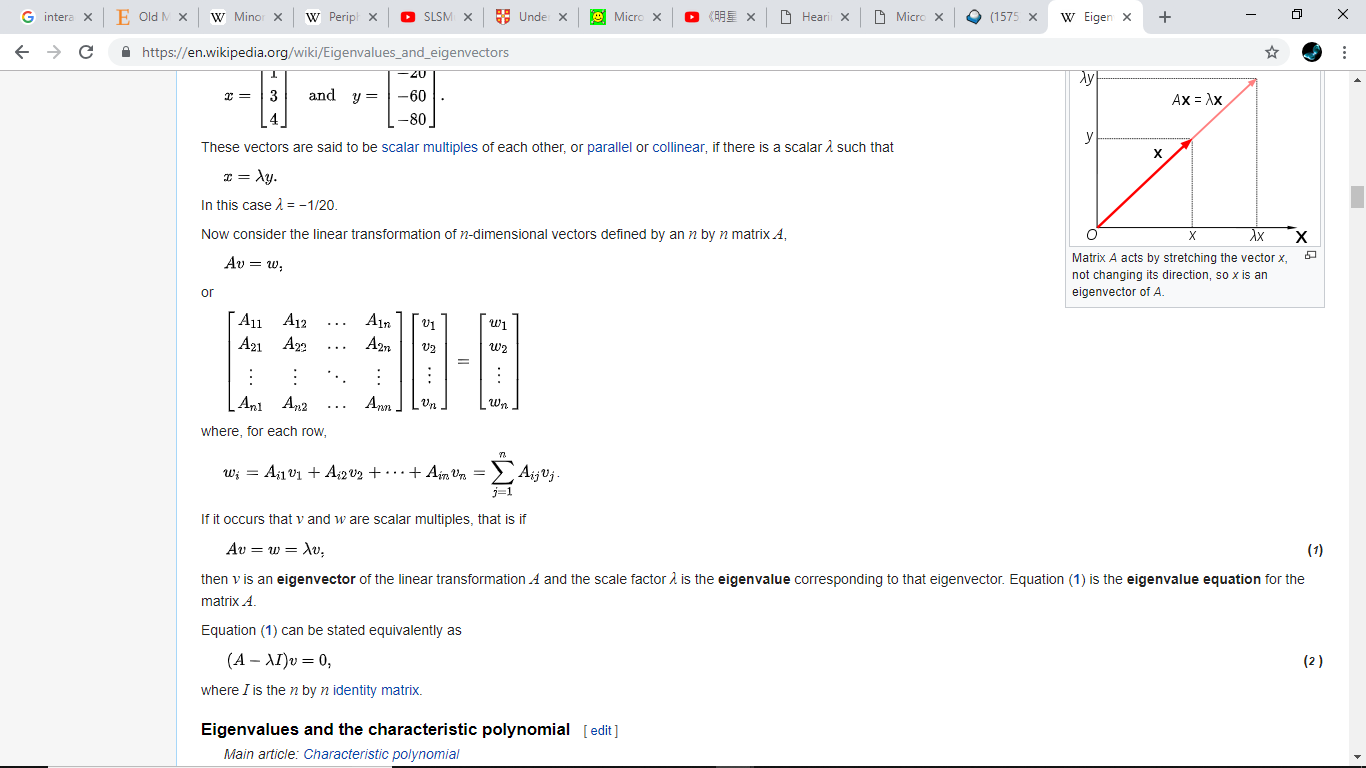
1. Fourier transform computes dot product btw sine waves of diff frequencies and EEG data
   1. Obscures time-varying changes in frequency structure of data
   2. Assumes data is stationary (mean, variance, frequency structure x change) – impossible for real daya. So perform temporally localised frequency decomposition methods (wavelet convolution, short time FFT) – assume stationary within brief periods
2. Morlet wavelet: sine wave windowed with a Gaussian
   1. How to make: sine wave multiplied by Gaussian. . . s = n/2πf (std dev).
   2. no sharp edges that produce artifacts, dampen influence of surrounding time points, control trade-off btw temporal precision and frequency precision
   3. constructing wavelet family: (1) frequencies cannot be slower than epochs; (2) frequencies cannot be above Nyquist freq; (3) 15 – 30 frequencies spanning 3 Hz to 60 Hz should suffice
   4. if results concern lower-frequency activity, use logarithmically spaced frequencies. If concerns higher-frequency activity, use linearly spaced frequencies
   5. wavelet should have same sampling rate as EEG data
   6. wavelets w less cycles more precise at localising dynamics in time (transient changes in activity, temporal precision); more cycles more precise at determining frequency of dynamic (temporally sustained activity; distinguish frequencies). Can increase number of cycles as frequencies increase (but data should be stationary during non-zero wavelet)
3. bandpass filtering + Hilbert transform
   1. Hilbert transform: compute Fourier transform > create a copy of Fourier coeffs and multiply by I > multiply +ve frequencies w -i (double original Fourier coeff), multiply -ve frequencies w I (subtract frm itself, becomes zero) > take inverse Fourier transform of modulated Fourier coeffs
   2. pros: more control over frequency characteristics of filter (can define inputs for bandpass filter shape; doesn’t have to be Gaussian). Cons: in Matlab toolbox, slower
   3. wider pass-bands (eg 20 Hz) for temporal precision; narrower pass-bands (eg 4 Hz) for frequency precision
4. short time FFT: use FFT to extract frequency structure of brief segments of data
   1. take small segment of data > taper ends > Fourier transform
   2. multitapers
      1. use when: (1) noisy data, small number of trials; (2) single-trial analysis, esp frequencies > 30Hz is impt; (3) focus on high-frequency power (> 60 Hz)
      2. do not use when: (1) frequencies < 30Hz; (2) if temporal precision is important

Other things to do

1. Normalisation (Power scaling): frequency spectrum shows decreasing power at increasing frequencies
   1. Decibel conversion . Baseline = period of time (few ms) before trial w no task-related processing, frequency-specific but not time-specific
   2. Percentage change and baseline division
   3. Z transform . Scaled to std dev units relative to power data during baseline period. Easily affected by variable data (noisy data, few trials)
2. Mean or median:
   1. Why not mean: no negatives for power ⭢ outliers contribute positive bias
   2. Use mean: clean data, sufficient no. of trials
3. Signal-to-noise ratio (SNR) = μ/σ at each point or after averaging
4. Intertrial phase clustering (ITPC): measures the extent to which a distribution of phase angles at each time-frequency-electrode point across trials is nonuniformly distributed in polar space
   1. Cannot take average of phase values (like power to form ERP) because phase angles are circular
   2. Method: compute average vector (not phase angles) and take length of average vector (0 < ITPC < 1). Further apart = smaller vector
   3. . Weighted . b is trial-varying vector
   4. Larger trial number = smaller ITPC. ITPC for frequencies < 15 Hz stabilise ard 20 trials, higher frequencies need more trials to stabilise

Spatial filters – data points set by weighted sum of activity at all electrodes

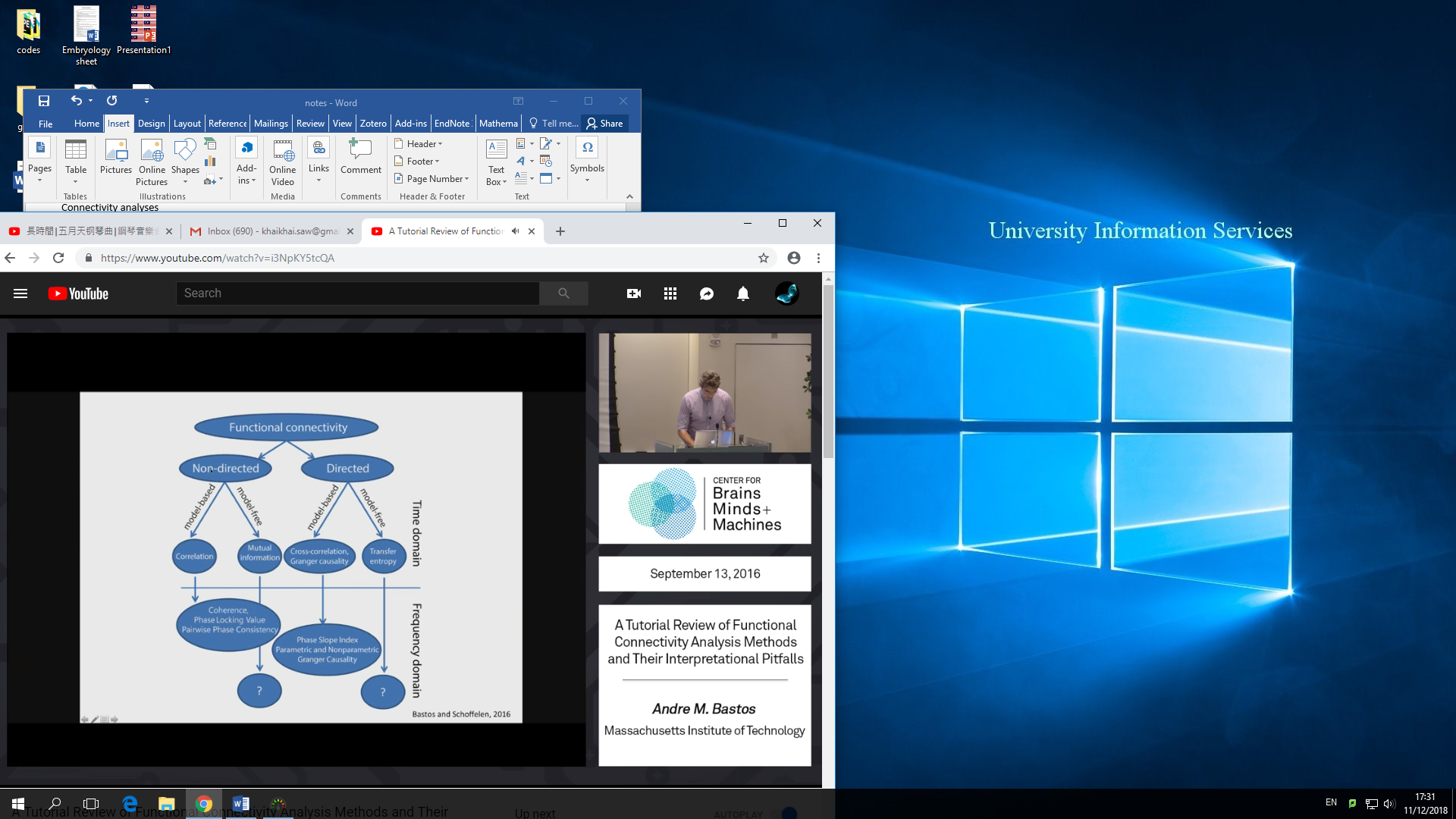
1. Surface Laplacian – interelectrode distances, filters out spatially broad topographical features
   1. Do not apply if results are from deep sources (insula) or by spatially distributed + highly temporally coherent generators. Must be applied on time-domain data (so surface Laplacian first before tf analysis)
   2. Method: weighted sum of activity at all electrodes is subtracted from activity of each electrode (first w G, then scaled w H) – second spatial derivative
   3. Electrode-by-electrode weighting matrices: . . M (2-6) – smoothness of result. P (10) – Legendre polynomial, high value allows only v high spatial frequencies
   4. . . .
2. Principal components analysis – construct weights (principal components) based on interelectrode covariances, based on statistical properties of data, physical locations not considered; highlights global features
   1. Variables (electrode) and instances (time points). Get covariance matrix from each trial then average tgt
   2. . X is electrode-by-time-points matrix. n – number of time points. Covariance matrix: diagonals contain variances of each electrode
   3. Eigendecomposition to return eigenvectors (new rotated axes) + eigenvalues (length of axes). Square matrix where nrow = nelec. Each column is principal component, each row stores weights of each electrode

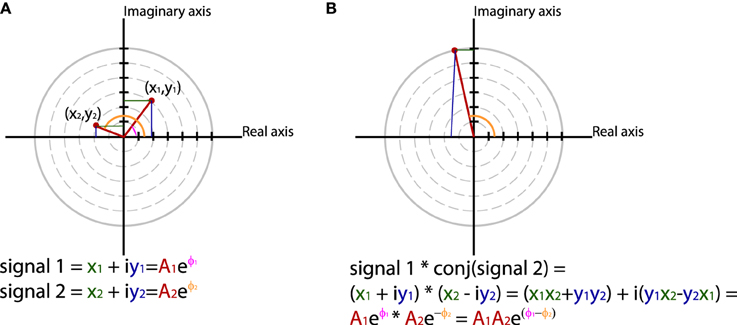
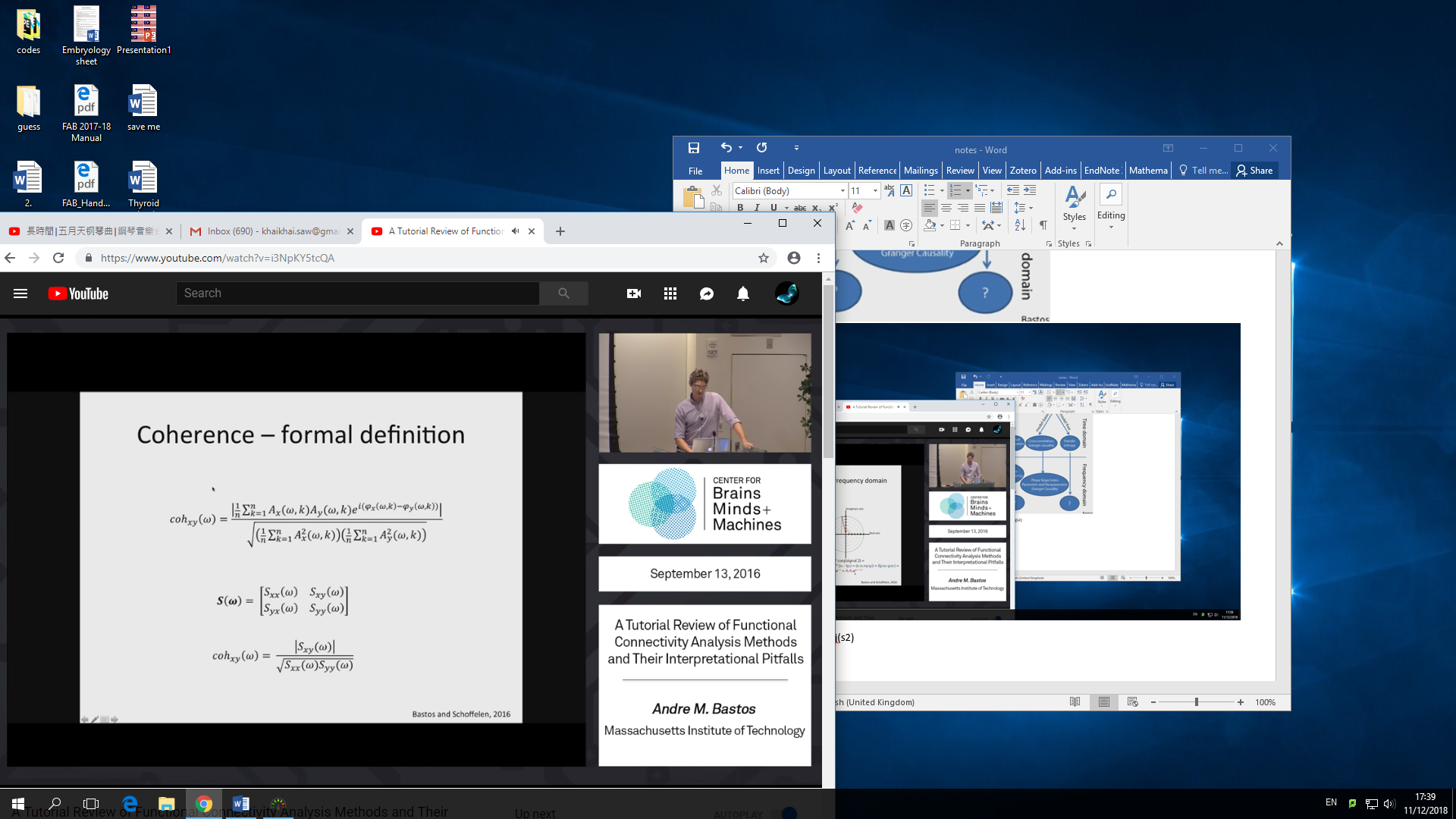
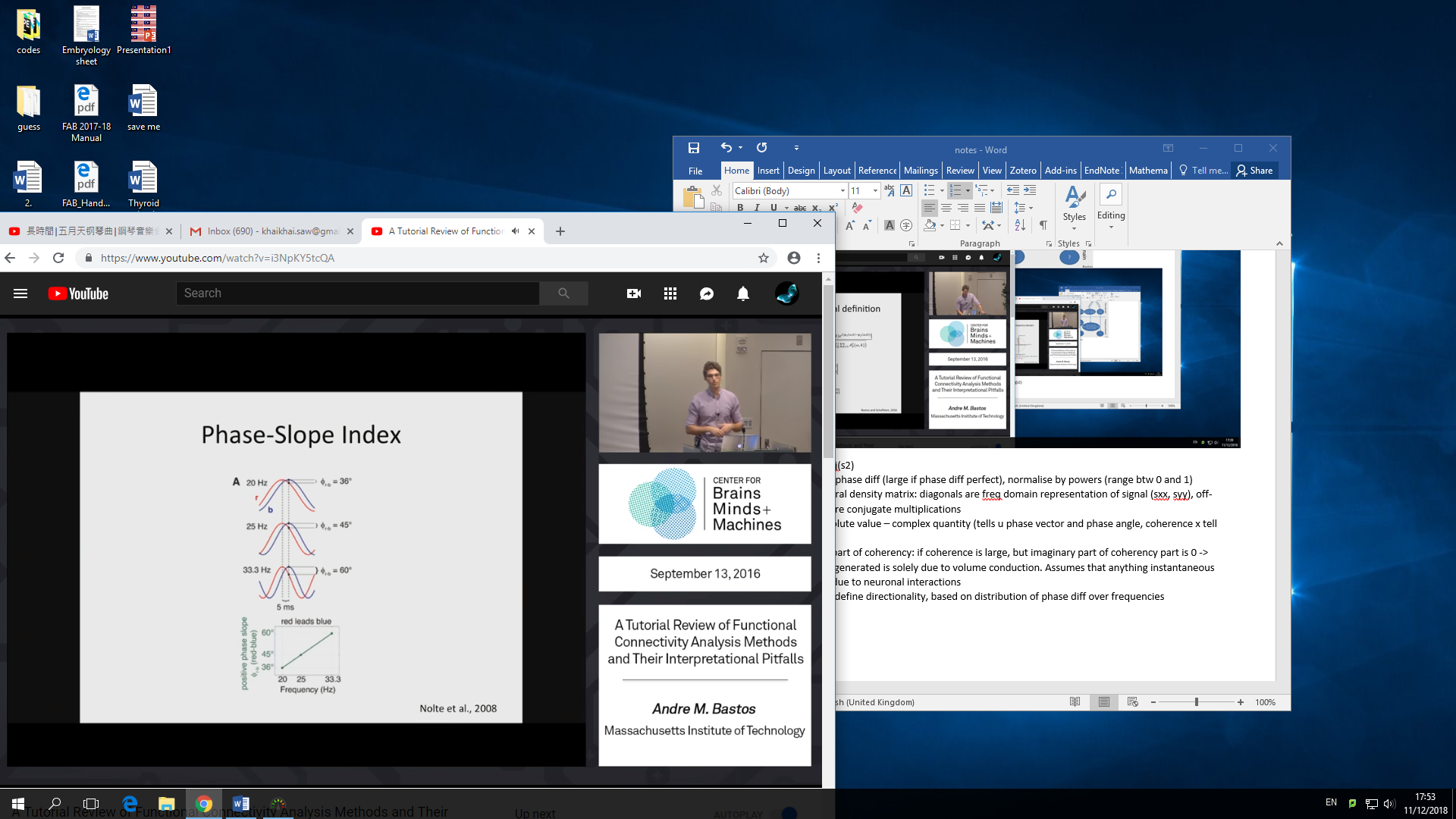
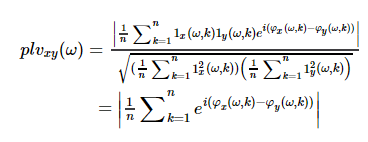


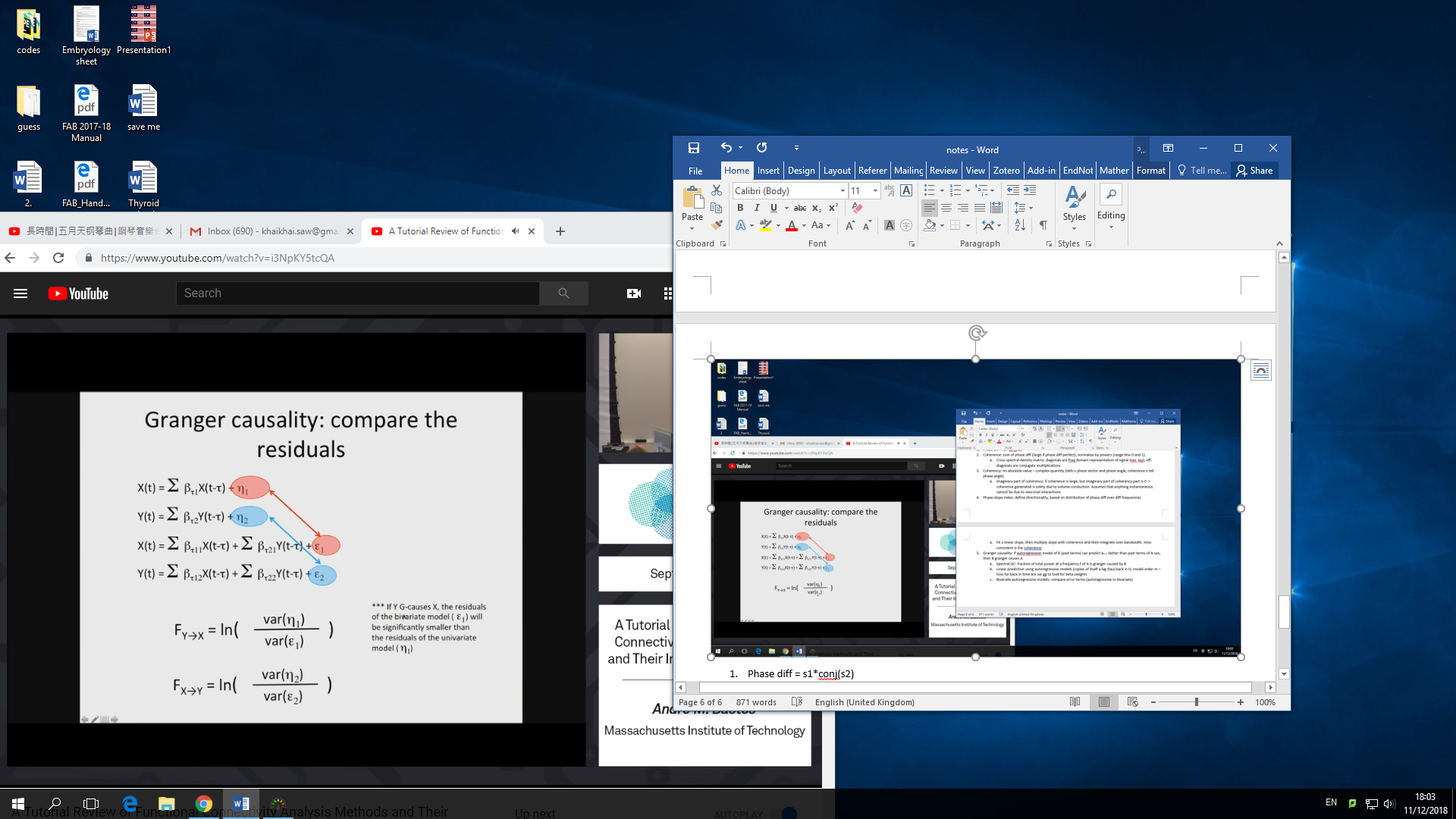
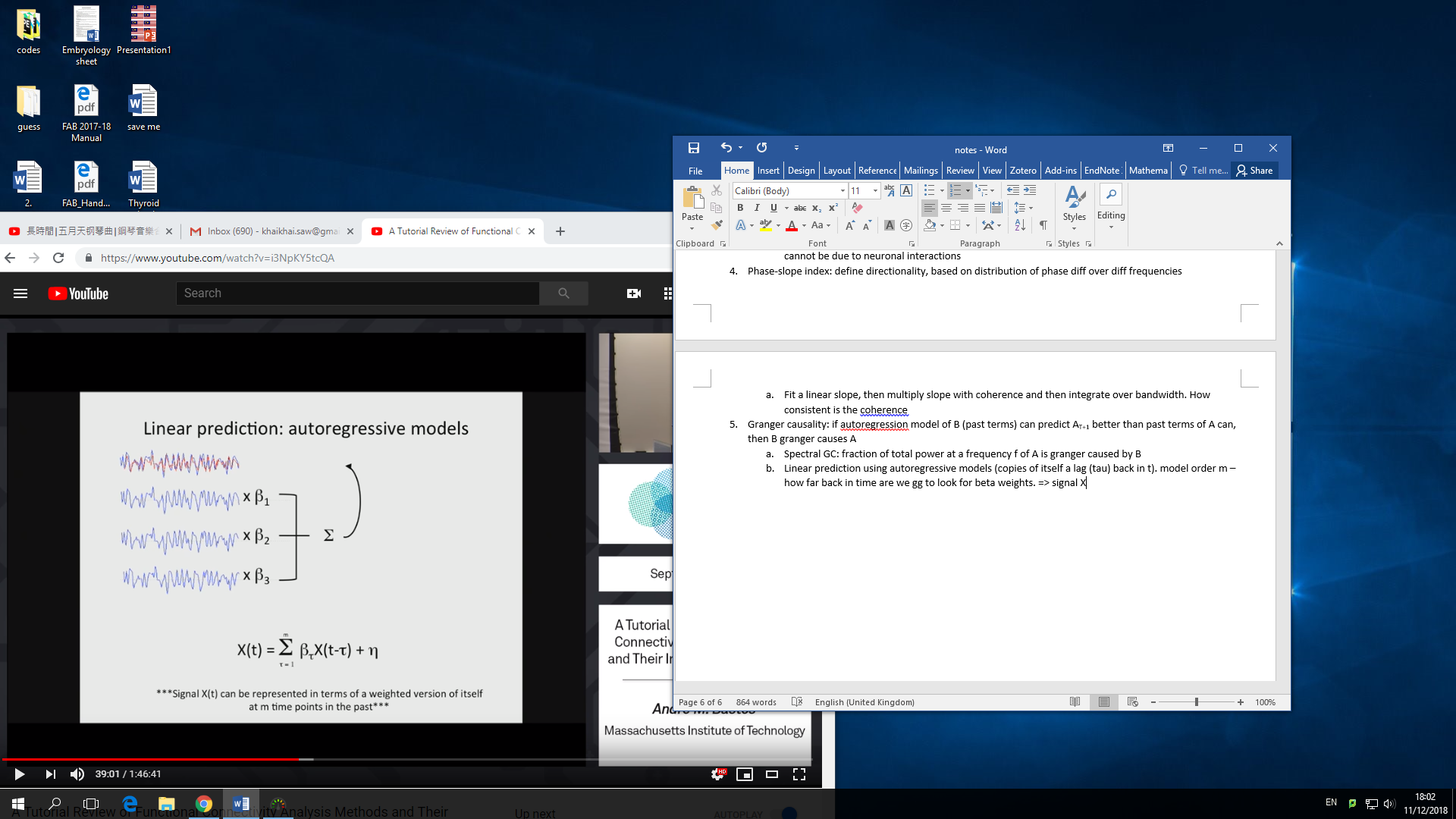
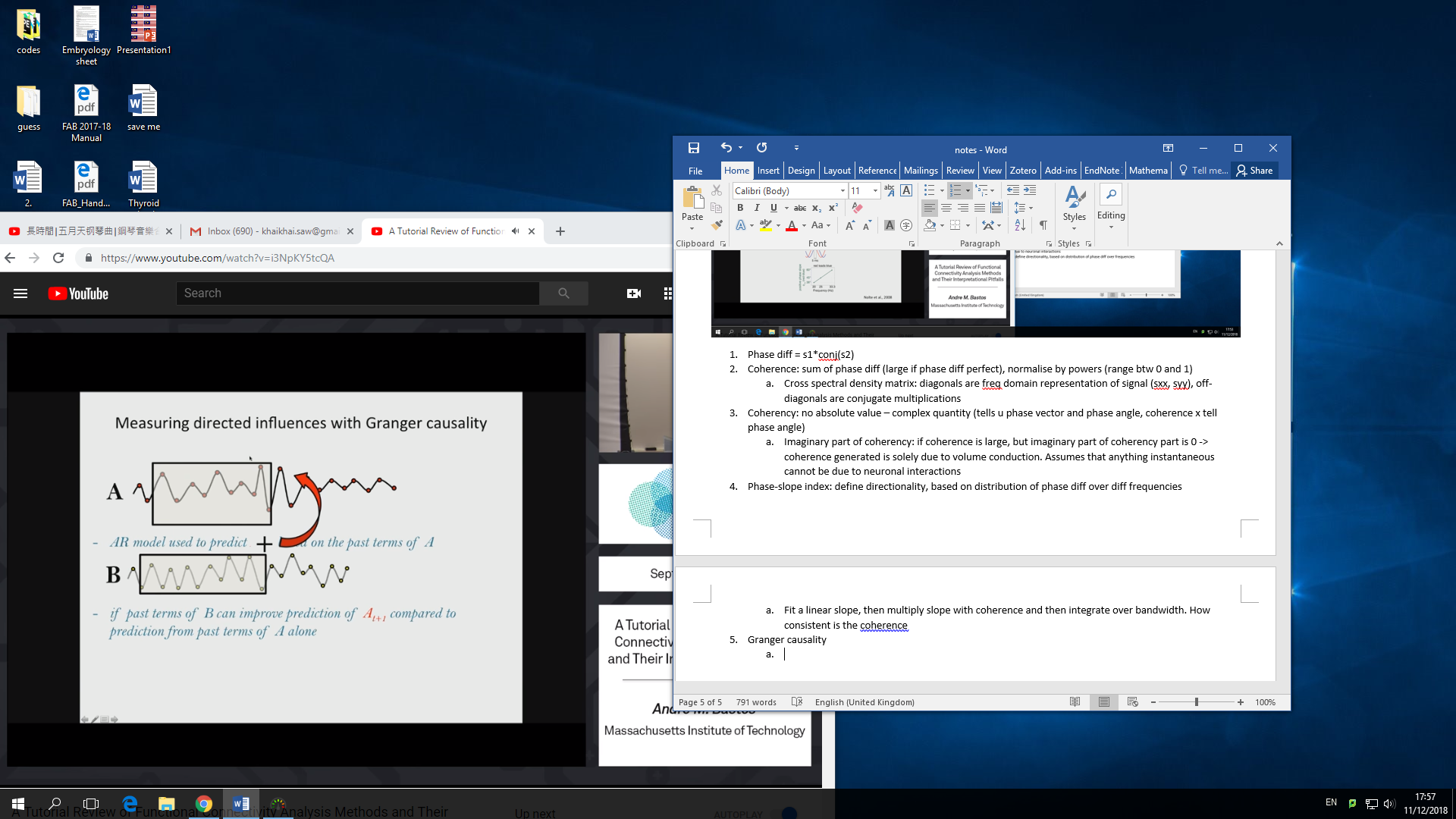
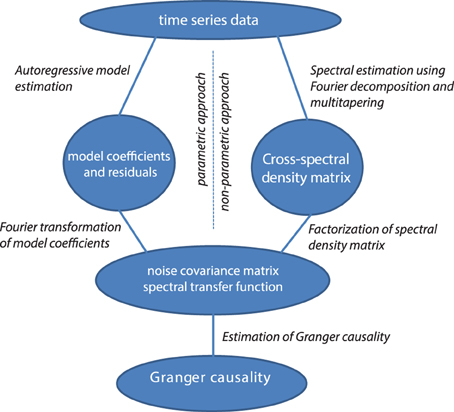
1. Beamforming – electrode activity + electrode position

Connectivity analyses

1. Concepts of bivariate connectivity
   1. Phase lag btw 2 electrodes x taken into consideration, as long as it’s consistent across time/trials
   2. Nonzero phase lag x mean causal or directed rship
   3. Phase reflects timing of activity within neural population – instantaneous connectivity; power reflects number of neurons/spatial extent of neural population – temporal offsets + jitters
   4. Functional connectivity (correlation) – linear/nonlinear covariation btw fluctuations in activity frm distinct neural networks. Effective connectivity (causation) – influence of activity in one neural network on another
   5. May be confounded by volume conduction (single signal detected bu multiple electrodes)
      1. If phase angle differences are strongly clustered ard 0, suggests volume conduction
2. Phase-based (see ITPC) – distribution of phase angle differences btw 2 electrodes (when neural popns are functionally coupled, timing of oscillation becomes synchronised)
   1. Cons: rely on precise temporal rships, susceptible to jitters (more significant at higher frequencies); no evidence for directionality
   2. Assumptions: connectivity is instantaneous, at same frequency
   3. Intersite phase clustering (ISPC) – clustering in polar space of phase angle differences btw electrodes. Instead of taking average of phase angles (ITPC), take average of phase angle differences btw electrodes over time. But non directional
   4. How to compute time-varying ISPC: (1) sliding time segments; (2) compute ISPC over trials instead of over time
      1. Pros of time: less sensitive to trial-to-trial jitter of expt event timing; insensitive to diff phase angle differences => better for high-frequency connectivity. Poor temporal precision => better for long task events. Can only be used if u have phase angle time series w same temporal resolution as ori data (complex wavelet convolution, filter-Hilbert but not stFFT, multitaper)
      2. Pros of trials: stronger evidence for task-related modulations in connectivity; no loss of temporal precision => time course of connectivity over tens to hundreds of ms. Sensitive to jitters => cannot perform on resting-state data. More sensitive to trial count
   5. Weighted ISPC-trials: correlate (Spearman’s correlation) ISPC value at each tf point over trials w trial-varying variable
   6. Spectral coherence – phase values weighted by power values
   7. Phase-lag based measurements: phase-based connectivity measure that ignore zero-phase lag (avoid spurious connectivity results from volume conduction)
      1. Imaginary coherence: same as spectral coherence, but imaginary part is taken before magnitude
      2. Phase-slope index: measures directed phase-based connectivity (whether slope of phase lag is consistently +ve/-ve over several adjacent frequency bins, sign of slope indicates directionality)
      3. Phase-lag index: measures extent to which a distribution of phase angle differences is distributed towards +ve/-ve sides of imaginary axis. If due to volume conduction, phase angle diff will be distributed ard 0 rad
      4. Weighted phase-lag index: weighted according to distance from real axis (vectors further away from 0/π rad have larger influence on estimate of connectivity
      5. Assumptions: phase lag + frequencies of activities of 2 electrodes are stationary during analysis
   8. When to use:
      1. ISPC + tests against volume conduction (prior hypothesis about small number of specific connectivity patterns)
      2. Phase-lag based measurements (data exploration) – resting state, tasks where connectivity strength is not compared across conditions
      3. Over trials (time course of changes, transient changes in connectivity) or over time (high-frequency, resting state, tasks w long event duration)
   9. Test mean phase angle – provides info about phase lag, directionality
      1. Gaussian v-test.
3. Power-based – can be btw same/diff freq and same/diff time points; essentially correlation (Pearson vs Spearman – no assumption of normal distribution); does not assume instantaneous connectivity
   1. Over time: get power time series (using time-frequency decomposition) ⭢ compute correlation coeff btw each power. Selecting length of time segment.
      1. Not limited to instantaneous connectivity (like phase based): perform cross-correlation analysis (reveal whether peal connectivity is observed when one time series is shifted)
   2. Over trials (3 methods): (1) specify time-frequency-electrode windows (hypothesis driven); (2) correlate power at each time points over trials (like ISPC-trials); (3) select time-frequency window from one ‘seed’ electrode and correlating w others (more open)
   3. Partial correlations: measure linear or monotonic rship btw 2 variables while holding a 3rd variable constant
      1. To test hypotheses abt networks comprising more than 2 nodes, minimise volume conduction artifacts during volume conduction
4. Granger prediction – test whether variance in one signal can be predicted from variance in another signal earlier in time; based on multivariate autoregression
   1. Can dissociate directional connectivity (A -> B or B -> A)
   2. Should not downsample (optimal sampling rates 250 – 1000 Hz)
   3. Selecting model order: Bayes information criterion (BIC, recommended for many data points like EEG) + Akaike information criterion
   4. Frequency domain Granger prediction: compute dot products btw autoregression coeff and complex sine waves, then apply those results to error variance via transfer function
   5. Granger prediction assumes stationary data. Make data stationary through: (1) detrending, (2) z-normalisation, (3) use shorter time segment
   6. Baseline normalisation
   7. Statistical analysis: (1) determine if prediction is significant (chi-square and F-statistics or permutation testing); (2) focus on differences in Granger prediction results over time/conditions/electrodes (nonparametric permutation testing)
5. Mutual information – detect shared info btw 2 variables; based on distributions of values within variables + joint distribution of 2 (or more) variables. Can detect many rships (linear or nonlinear rships that correlation cannot detect). Can tell relationship (but don’t know positive or negative)
   1. Shannon entropy (from information theory): amt of information or ‘surprise’ a variable has
   2. Compute entropy: bin data (create histogram) ⭢ compute probability that a value of data would fall in each bin (bin count/total bins) ⭢ multiply probability value by log2(probability value) ⭢ sum and multiply by -1
   3. No subject to power-law scaling effects, don’t need baseline normalisation; unrelated to temporal structure of data. Higher entropy means system can take more states or configurations. Joint entropy = total entropy of a pair of variables – involves codistribution of two variables
   4. Mutual information: entropy 1 + entropy 2 – joint entropy
   5. Small trial numbers will inflate estimates. Noisy data cannot be attenuated by having more noisy data. SNR of frequency-band-specific power higher than broadband EEG. If computing mutual information on time-domain data, should low-pass filter EEG data to remove high-frequency fluctuations + noise
   6. Lagged mutual information: time shift one signal repeatedly for multiple time lags
   7. Statistics: (1) use nonparametric permutation testing to transform bits to standard statistical Z value
6. Cross-frequency coupling – statistical rship btw activities in 2 freq bands; infer local organisations (measured at single electrode) + long-range connectivity (diff electrodes)
   1. Power-power correlations: (1) correlate 2 power time series over time; (2) cross-trial time-frequency power correlations
   2. Phase-amplitude coupling (most common): (1) power and phase over time; (2) power in phase space; (3) power distribution in phase bins
      1. Phase-amplitude coupling (PAC) measured by length of average vector. To prevent influence by power values: (1) scale by power; (2) convert power values to rank and apply Moore’s modification pf Rayleigh test). Nonuniform distribution of phase angles, large outliers might also affect. So apply nonparametric permutation testing to solve all 3 confounds
      2. To make sure it is not task-related phase and power coactivation (diff frequencies reacting to stimulus independently): (1) plot ITPC from lower frequency to show PACZ and ITPC do not temporally co-occur. (2) avoid analysis during time periods w strong ITPC. (3) subtract ERP from single trial EEG data to eliminate ITPC
      3. Exploratory: set one frequency but the other unsure. Do for broad frequency range (eg 2-15Hz for phase, 20-100 Hz for power)
      4. Notes: (1) best done using wavelet or filter-Herbert (returns phase angles), (2) use wavelets w small number of cycles (3-5) or filter-Hilbert w wide freq bands (timing is impt), (3) there is ⭣ SNR at higher freqs
   3. Phase-phase coupling
7. Pairwise spike-time coupling – treat as 1 or 0 (I think..)
8. Graph theory – multivariate networks; nodes/vertices (electrodes) and edges (connectivity)

A Tutorial Review of Functional Connectivity Analysis Methods and Their Interpretational Pitfalls

1. Overview
   1. Non-directed: model-based (assumes linearity – eg. Pearson correlation coeff) vs model-free (eg. mutual information). Ignores temporal structure, but can obtain cross-correlation function to study directionality
   2. Directed: model-based (eg. Granger causality – linear autoregressive model or non-parametric spectral matrix factorisation) vs model-free (eg. transfer entropy)
2. Usually study signals in frequency domain: transformation achieved through non-parametric (Fourier decomposition, wavelet analysis, filter-Hilbert) or parametric (autoregressive models)
3. Measures of synchronisation:
   1. Represent amplitude + phase as *Aeiϕ*or *x + iy* ⭢ combine spectral representation of individual signals to obtain cross-spectral density (frequency domain equivalent of cross-covariance function) ⭢ frequency-wise multiplication s1\*conj(s2) to give complex number (magnitude – product of s1 + s2 amplitudes; angle – phase diff)
   2. Then take weighted sum of cross-spectral densities (draw all vectors head to tail and normalise) ⭢ if there is consistency in phase diff, length of weighted sum will have non-zero value
4. Coherence: frequency domain equivalent to time domain cross-correlation function
   1. sum of phase diff (large if phase diff perfect), normalise by powers (range btw 0 and 1)
   2. Cross spectral density matrix: diagonals are freq domain representation of signal (sxx, syy), off-diagonals are conjugate multiplications
   3. Coherency: when |…| is omitted from coherence – phase diff angle may be interpretable
5. Imaginary part of coherency: if coherence is large, but imaginary part of coherency part is 0 -> coherence generated is solely due to volume conduction. Assumes that anything instantaneous cannot be due to neuronal interactions
6. Phase-slope index: define directionality, based on distribution of phase diff over diff frequencies
   1. Fit a linear slope, then multiply slope with coherence and then integrate over bandwidth. How consistent is the coherence. If there is consistent phase diff changes, PSI will deviate from 0 (sign informs directionality)
   2. Problem: If A leads B and B leads A, then results will be 0 directionality
7. Phase locking value (PLV): apply coherence eqn to amplitude normalised Fourier transformed signals.
   1. Use this instead of coherence because less amplitude correlations
   2. 
8. Phase lag index (PLI) – weighted, debiased weighted: evaluates distribution of phase diff
   1. Computed by averaging sign of estimated phase diff
9. Pairwise phase consistency (PPC): quantifies distribution of phase diff
   1. Computed from distribution of all pairwise differences of relative phase (not vector average like PLV)
   2. Not biased by sample size like PLV
10. Granger causality: quantify bidirectional interactions (both ways)
    1. If autoregression model of B (past terms) can predict AT+1 better than past terms of A can, then B granger causes A
    2. Spectral GC: fraction of total power at a frequency f of A is granger caused by B
    3. Linear prediction using autoregressive models (copies of itself a lag (tau) back in t). model order m – how far back in time are we gg to look for beta weights
    4. Bivariate autoregressive models: compare error terms (autoregressive vs bivariate). If error terms of bivariate is much smaller, then granger causality)
    5. Parametric (autoregressive model based -> Fourier transform, assumes residual is white noise) vs nonparametric (FFT -> factorise into directed [spectral transfer function] and non-directed components [noise covariance]
       1. Advantages of nonparametric: dunnid to define autoregressive model order

1. Bivariate vs multivariate

Spike trains and firing rate

1. Bin (count nspikes in specified time frame) ⭢ rectangular/Gaussian window function