1. Lecture 3.1 – The role of Signal Processing
   1. BCI Theory
      1. BCI leverages theory from a wide range of fields
         1. Signal Processing, Machine Learning, Statistics, Neuroscience, Control Theory, Information Theory
      2. A given BCI may be understood from the vantage point of any of these theories
      3. But no signal theory conveniently describes all aspects of a BCI
   2. Signal Processing - Types
      1. Digital Signal Processing
         1. Concerned with systems (a.k.a. filters) that transform one signal into another
      2. Linear Time-Invariant (LTI) Systems
         1. Spectral filters and their optimal design
         2. One of the most developed areas (best developed in the field of signal processing)
      3. Statistical Signal Processing and Adaptive Filtering
         1. Among the advanced areas
            1. Kalman Filter (well known)
            2. Recursive least squares, etc.
      4. Sparse Signal Processing
         1. Sparse recovery and compressive sensing
         2. New branch with application to BCIs
         3. Last ten years or so
         4. The most sophisticated in BCI you can really do
   3. Signal Processing – Diving In
      1. A signal is a mapping from an index set (here discrete time) onto vectors (multichannel samples)
      2. From the point of view of Signal Processing, a BCI transduces the input signal (for example EEG) into a control signal
      3. It is defined by a transformation rule T
   4. Important System Types
      1. Static vs. Dynamic
         1. A system is called ***static*** if the value at any sample n depends only on , otherwise ***dynamic***.
      2. Causal vs. Non-Causal
         1. A system is ***causal*** if the output at any time n only depends on values of for , otherwise ***non-causal***.
         2. It does not look into the future
         3. All BCIs are non causal because it cannot look into the future
      3. Time-Invariant vs. Time-Variant
         1. A system is called ***time-invariant*** if implies that for every time shift *k*, otherwise ***time-variant***.
         2. If you shift time, does this shift the output….
      4. Linear vs. Nonlinear
         1. A system is called ***linear*** if the equation holds for all inputs and and all constants and , otherwise ***nonlinear***.
         2. Twice the input signal gives you twice the output signal
   5. BCIs Viewed as Filters
      1. Since BCIs are operated in real time, they are always *causal* systems
         1. Any of its part have to be causal
      2. BCIs usually perform temporal filtering, and are therefore *dynamic*
         1. They do transfer information across time
      3. Some BCIs are *time-invariant*, but adaptive BCIs are not
      4. Simple BCIs are *linear*, but the vast majority is not
      5. Since their output is needed at a much lower sampling rate (0.1-60 Hz) than the input (250-1000 Hz), they are technically *multi-rate* systems
   6. BCI Components as Filters
      1. BCI components are conveniently described as filters – more so than the entire system itself
      2. This gives rise to several key categories of filter components
2. Lecture 3.2 Major Filter Classes
   1. Static Filters
      1. Signal Squaring
         1. Static system useful step in calculating the variance of the signal
      2. Logarithm
         1. Useful later…
   2. Spatial Filters
      1. Operate across space, not time
      2. Transform a multi-channel signal such that each depends only on ; most spatial filters are linear, i.e. for some matrix
         1. Once you have calculated the matrix M you can apply is linearly throughout the whole system
      3. Linear spatial filters can approximately invert volume conduction and remap channel signals to approximate source signal – this is their main use in BCIs
         1. The mappings from a channel are linear and can use them to trace the source
         2. Don’t want to operate at the level the sensor, you want to operate at the source
         3. Subtract the average of all channels from one channel
         4. Subtract the average of all neighbors…
      4. Examples:
         1. Re-referencing, Surface Laplacian, Independent Component Analysis (ICA), Common Spatial Patterns (CSP) – more later
      5. Spatial Filters Visualized
         1. Spatial filters designed to recover motor-cortex source activity, calculated via the Common Spatial Patterns algorithm
   3. Spatial Filters vs. Forward Projections
      1. Spatial filters are not the same as forward projection maps of some source signal
      2. They are the inverse operation
      3. When ever you see a visualization of a small area in the brain like a source projection, it is a spatial filter
         1. The new signal is some matrix W times the signals X(n)
      4. If you see a dipole like visualization it is most likely a forward projection map
   4. Temporal Filters
      1. Transform a multi-channel signal such that each channel in depends only on the channel
         1. They do not go across channels so they do not go across space, they transfer data across time
      2. They are conceptually orthogonal to spatial filters
      3. Examples include time windowing, wavelet transform, etc.
         1. Setting time equal to zero
      4. Special case: Spectral filters
      5. Moving Average  
         1. The *i*th channel of the input is equal to the sum of n samples of the input divided by m samples
         2. Effectively a smoothing (low-pass) operator
         3. In fact a simple example of a spectral filter
         4. Also called running mean, removes high frequency noise
   5. Spectral Filters
      1. Temporal filters that are designed for their effects on the spectrum of the signal
         1. Designed or optimized to have a particular effect on the frequency components of the signal
      2. Spectrum of a signal
         1. A representation of the signal as a sum of N sinusoidal components,  
            where is the amplitude of each sinusoid and is its phase
         2. Every signal is simply the sum of all signals made up of some sin wave defined by the phase offset multiplied by some amplitude…
      3. An equivalent (more common) representation is the Fourier Series representation  
         where is not complex-valued and represents both the amplitude and the phase
      4. This relies on the Euler (oiler) formula
      5. Examples include
         1. High-pass, low-pass, band-pass, notch filters
      6. Their main utility in BCIs is to isolate oscillations or ERPs of interest
      7. FIR (Finite Impulse Response) Filter  
         1. A key spectral filter
         2. Performs convolution between signal and kernel
         3. The trick lies in the coefficients (“kernel”)
         4. Can implement any linear time-invariant spectral filter
            1. The sum of some coefficients times some part of the filter
         5. Moving average is a special case
         6. Allows you to have a variable coefficient for time
      8. FIR Filter Design in MATLAB
         1. Various criteria for filter kernel design (given a desired frequency response) (in the signal processing toolbox)
            1. Least-squares error: firls
            2. Minimax error (Parks-McClellan): firpm
            3. Using the Fourier transform: fir2
         2. Choice of a reasonable filter order (length):
            1. firpmord
         3. Minimum-phase filter design (using Cepstral analysis): rceps
   6. Other Filters
      1. Spatio-Temporal Filters are also used, but have too many degrees of freedom to be hand-designed
         1. Usually the result of an adaptive procedure
         2. Apply different frequency weightings for different parts of the brain
         3. Need lots of data (knowledge) to design these
      2. Spectral Transforms, which transform between the time and spectral representation of a signal are used frequently as intermediate stages
         1. Take a time domain signal and transform it to the frequency domain
         2. Linear transform
      3. Rate-changing filters such as resampling are useful to manage computational costs
         1. Takes a high rate signal and transforms it to a low rate signal.
3. Lecture 3.3 – A Simple Neurofeedback BCI
   1. A Simple Neurofeedback BCI
      1. Feed back the amplitude of a brain idle oscillation (e.g. 10 Hz alpha for relaxation) to the user/subject
         1. A person wants to train themselves to be relaxed
      2. Also other processes conceivable
      3. Can be implemented using discussed tools:
4. Lecture 3.4 Prediction Function Notation
   1. Alternative to the Signal Processing Framework
      1. A BCI with a limited memory of the past could also be viewed as a mathematical mapping :  
         ; someSignal then   
         “subj. excited” (+1),   
         “subj. not excited” (-1)
         1. Use this with an signal processing
         2. X is a matrix of signals
         3. y can be binary or anything else…
      2. The functional form is arbitrary, for example
      3. The mapping involves unknown parameters, here **W** and b
   2. Functional Form
      1. Reflects the relationship between observation (data segment X) and desired output (cognitive state parameter y)
      2. Based on some assumed generative mechanism (forward model) – or ad hoc
         1. Can try to find a location to monitor
      3. Note: Functional form is the inverse mapping
   3. Core Ingredient: Spatial Filter
      1. Linear inverse of volume conduction effect  
          (forward)  
          (inverse)
      2. The observation (chunk of EEG) is some matrix times the equivalent number of source samples would be the forward mapping
      3. If that mapping is invertible then we can get the source signal from just the signals
   4. Full Examples (Details Later…)
      1. Inverse mapping from filtered source time course to latent cognitive state, e.g.:  
          (linear)  
          (nonlinear…)
      2. Still have to deal with how the source time courses, say a ripple, in a bunch of neurons relates to the cognitive process (“thought”) or parameter that we are trying to infer
   5. Neurofeedback BCI in Functional Style
      1. Performed as a mapping of a sliding window **X** onto the output y:  
         1. **X** is the chunk of signals
         2. **T** is a linear of transform that is number of samples by number of samples matrix
      2. **T** implements a temporal filter, written as a matrix multiplication (each column = shifted filter kernel)
   6. In Comparison…
      1. Main drawback of the pure mathematical form compared to the signal processing approach:
         1. The entire input window X is re-processed for each desired output value y
         2. Especially bad if the window X moves only by a few samples between evaluations of f
         3. In contrast, most signal processing methods are incremental of recursive
            1. FIR or IIR
         4. If you apply this function to a chunk of EEG and then 10 seconds later you apply to another chunk of EEG, you are doing a lot of redundant processing.
         5. If you had used signal processing to implement you would have used filters that have been tuned for many years, where a filter may use components from the previous computations
      2. Main benefit is the relative conceptual simplicity
   7. In Combination
      1. Both frameworks are complementary, rather than contradictory, and are in practice often used in combination
      2. Prediction function is queried on demand
      3. i.e. we filter and then apply out expensive prediction function periodically
   8. Neurofeedback Using the Combined Approach
      1. Computationally costly spectral filtering is done in the signal processing portion
      2. Lightweight predictive mapping is done at the lower rate in the function portion
   9. Outside View
      1. Implemented this way, BCIs act as an oracle that consumers one or more multi-channel signals and can respond to queries about a pre-defined question
      2. Note: in modern BCIs the output is often a discrete probability distribution
         1. Not a 1 of 0 output