# Data-driven Connectivity Changes in Patients with Alzheimer's Disease

Andrew Kerr, Catherine Chesnutt, Kwaku Akrofi, Mary C. Baker
Texas Tech University
Autumn's Dawn Neuro-Imaging, Cognition, and Engineering Laboratory

## Introduction

The goal of this project was to apply data driven methods to identify differences in EEG patterns between patients with Alzheimer's Disease (AD) and control subjects. Features were optimally selected that best separated AD patients from age-matched controls using sequential forward floating selection (SFFS). Some features may be useful in classification within a group, but may not show individual statistical significance when using statistical tests like the t-test or the Wilcoxin test. The feature sets are validated based on classification rates, and the classification rates from features selected by SFFS are compared to the classification rates of features passing the t-test. The working hypothesis is that coherence features of EEG patterns are the most effective for separating patients with AD from control subjects and can be used to model those areas in the brain which experience a change in connectivity with the onset of AD.

## Methods

Resting EEG was recorded on 32 subjects: 16 AD patients and 16 control subjects. Each subject's EEG was recorded for 30 seconds using a clinical Medelec-Valor system with a 10-20 referential montage consisting of 19 channels. Patients were selected from the Memory Disorders Clinic, and were diagnosed as having AD using physical and neuropsychological exam results and the DSM IV criteria for Alzheimer's Disease. Controls were typically age-matched spouses. The data were normalized by taking the RMS of each channel for each subject and dividing each sample from the channel by the RMS value for the channel.

The EEG average band power was determined for each channel for each of the 5 common frequency bands: delta, theta, alpha, beta, and gamma. EEG coherence was determined for the channel pairs described in [1] and shown graphically in figure 2. These sets of channel pairs correspond to known connectivity paths in different brain regions. The coherence of all possible channel pairs was calculated to compare with the sets of channel pairs known to be significant.

A plug-in was written for EEGLAB[2] that applies the appropriate frequency band filters, generates the features, and selects those features which are optimal for classification. A sequential forward floating selection (SFFS) [3] algorithm determines which features result in effective separation of the AD from control groups. The unique aspect of the SFFS algorithm is that it tests classification rates for *groups* of features, rather than individual features, to find the best feature set for classification. The block diagram of the SFFS algorithm is shown in Figure 1. A linear discriminate analysis classification technique along with the leave-one-out method of cross-validation are used to validate sets of features. The maximum number of features selected is set to four since the number of subjects is 32.

The p-value of each generated feature is calculated from the t-test and the classification rate is assessed using the four features with the lowest p-value. The comparison of these two methods are represented in table I.
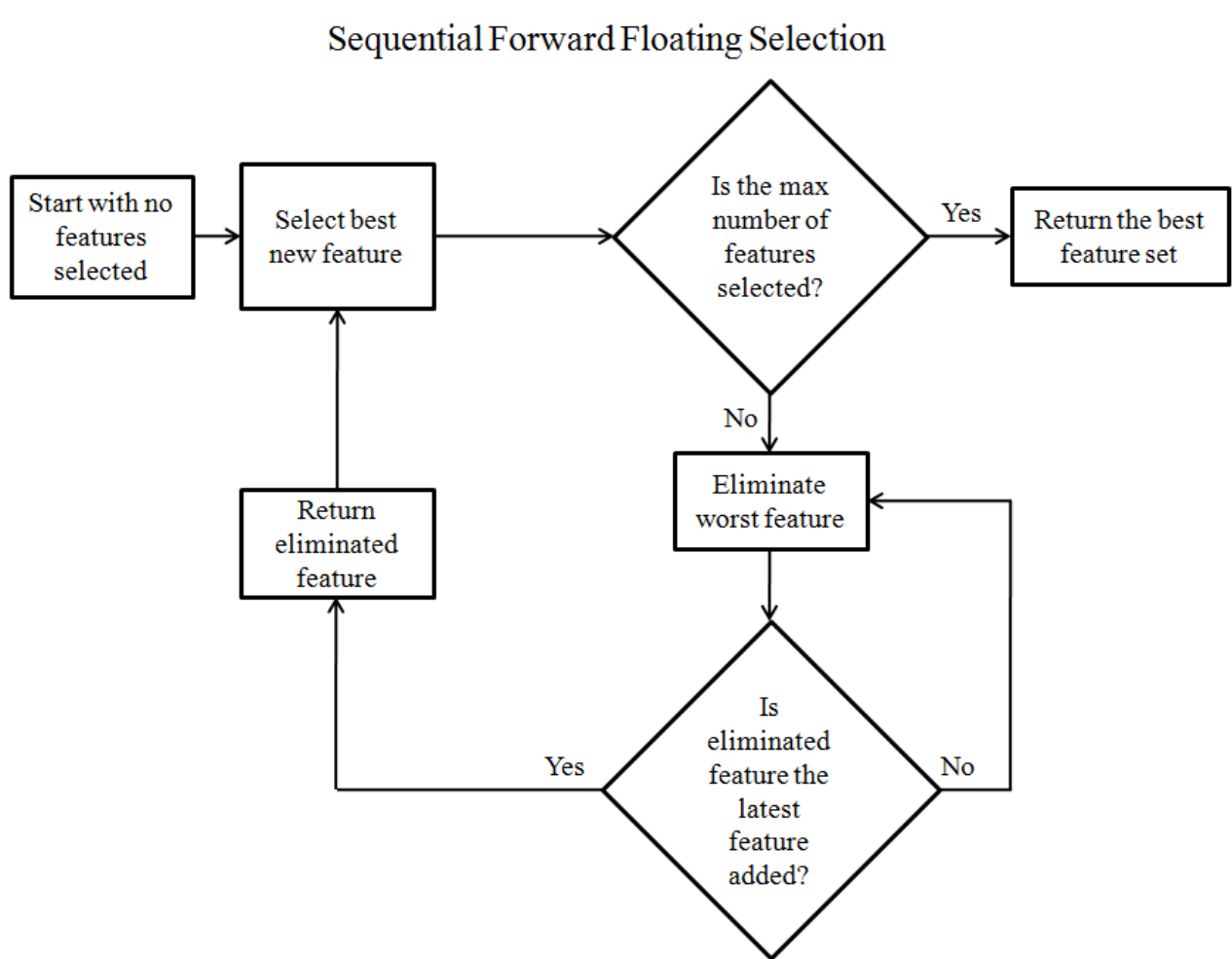


Figure 1 – Sequential forward floating selection block diagram.

Table I – Comparison of the classification rates of features selected by SFFS, with a maximum of four features, and the four features with the smallest p-value from the t-test, for the difference channel pair combinations.
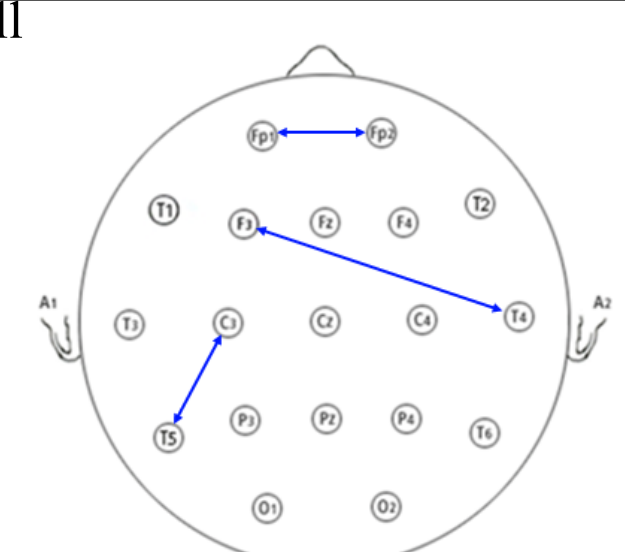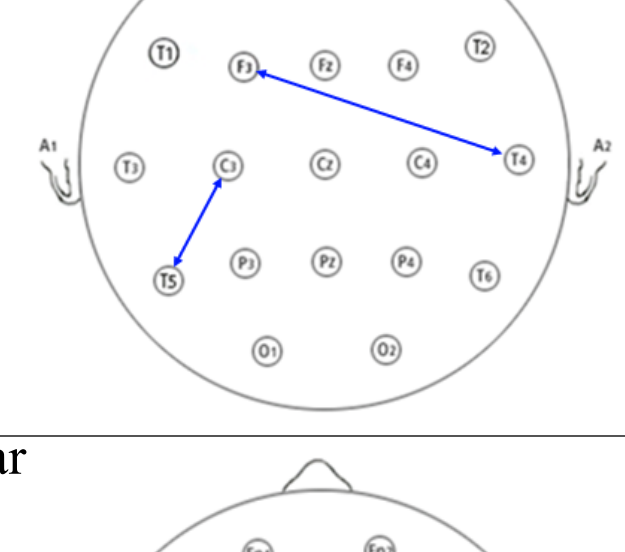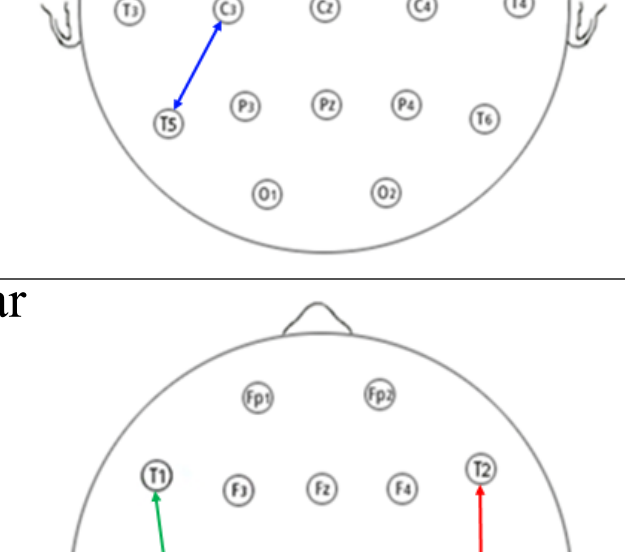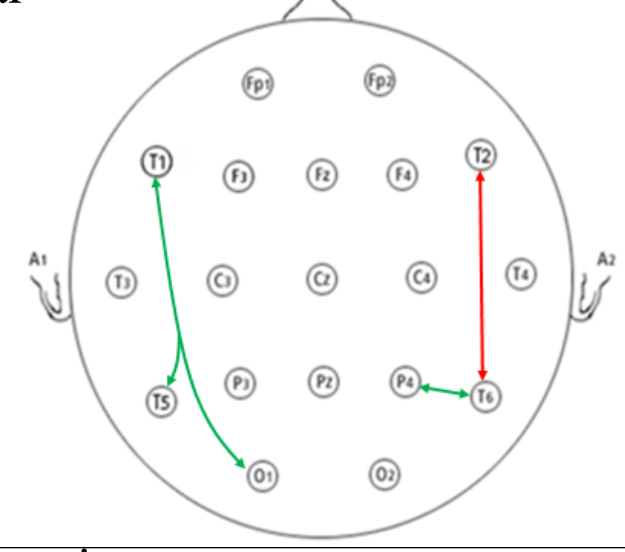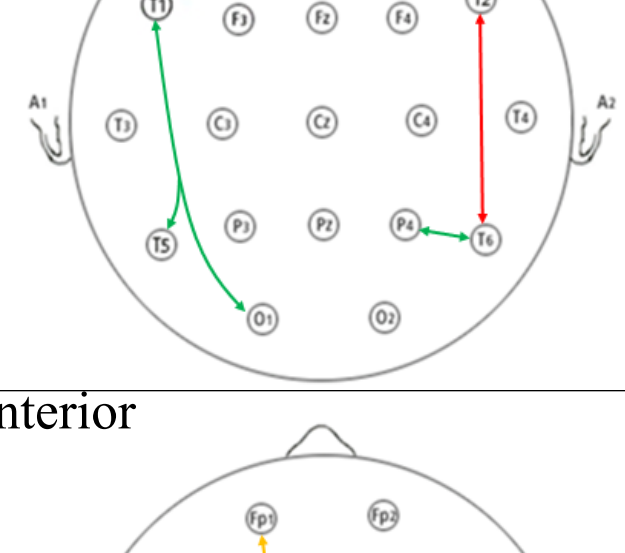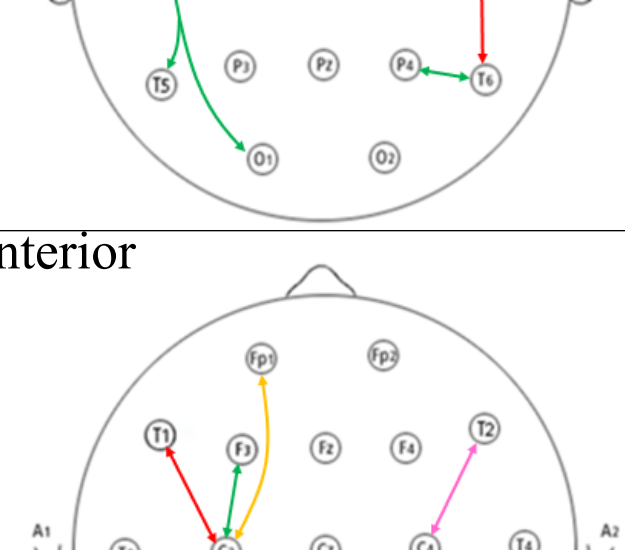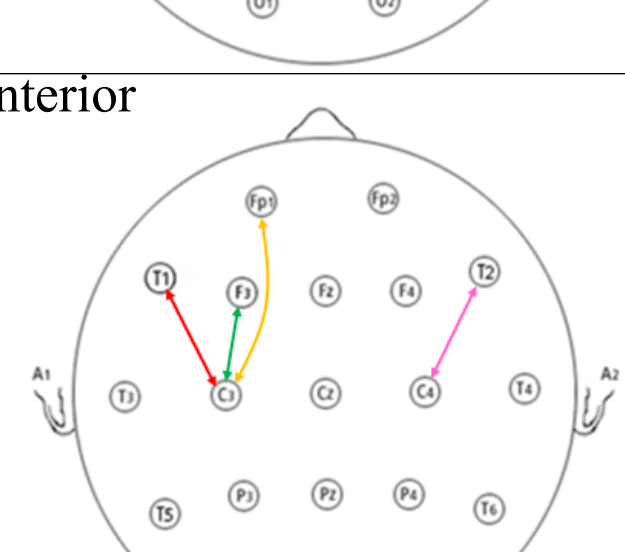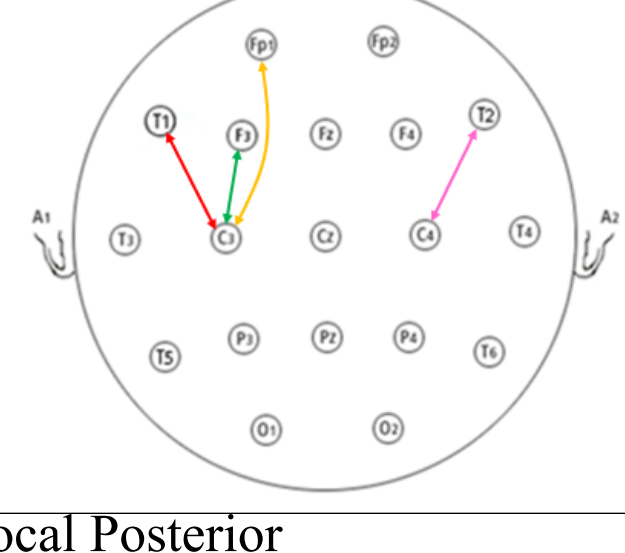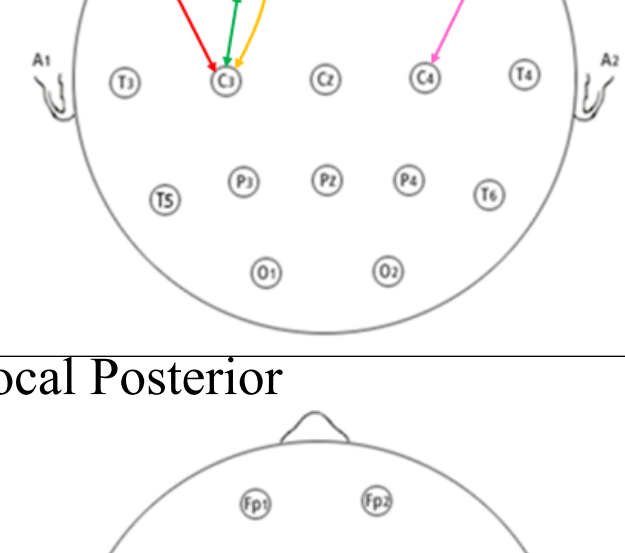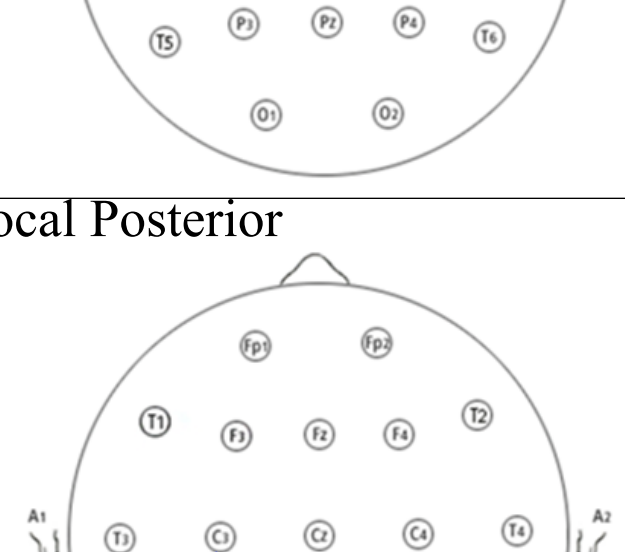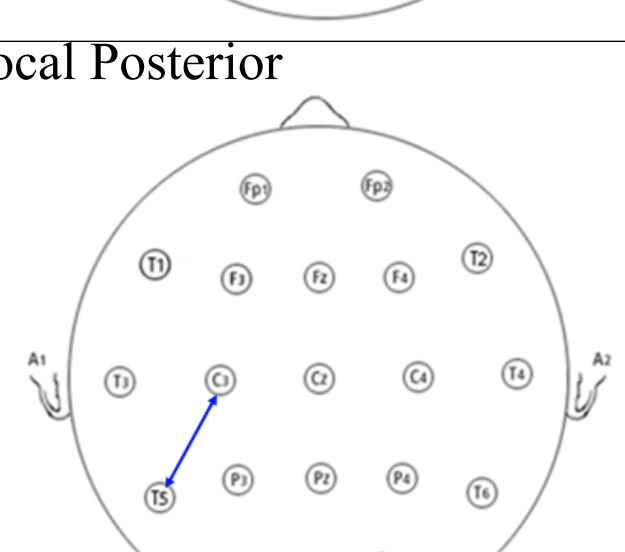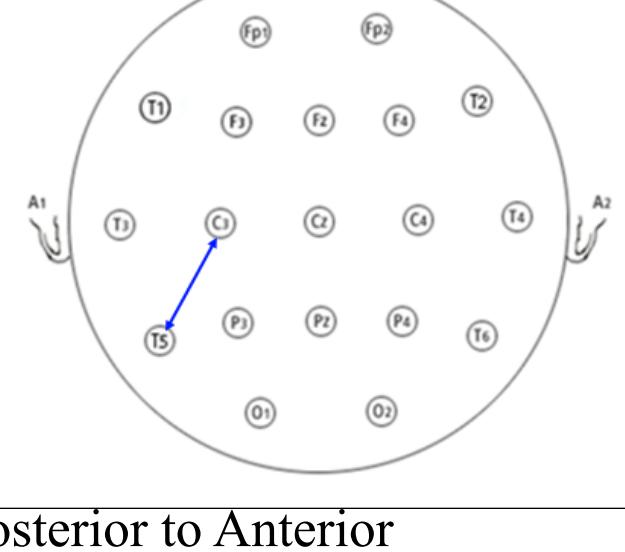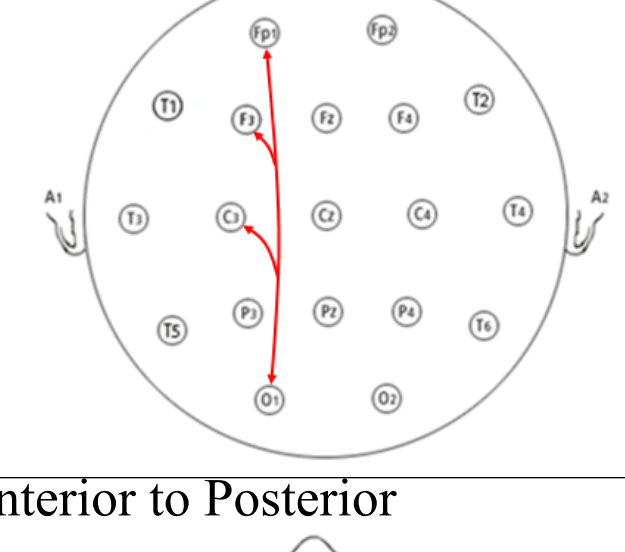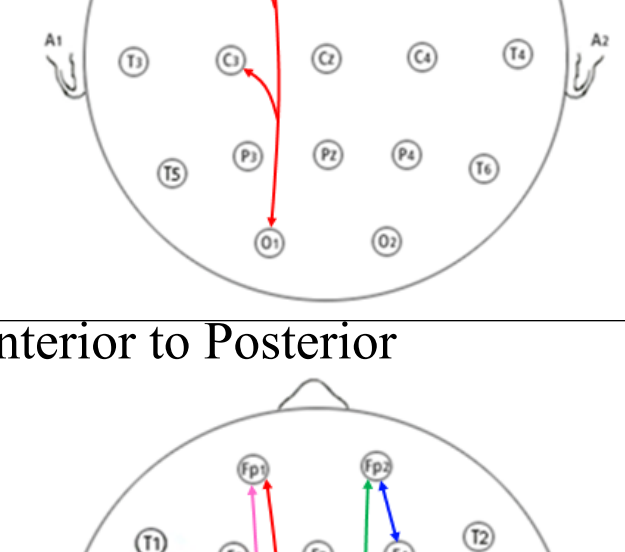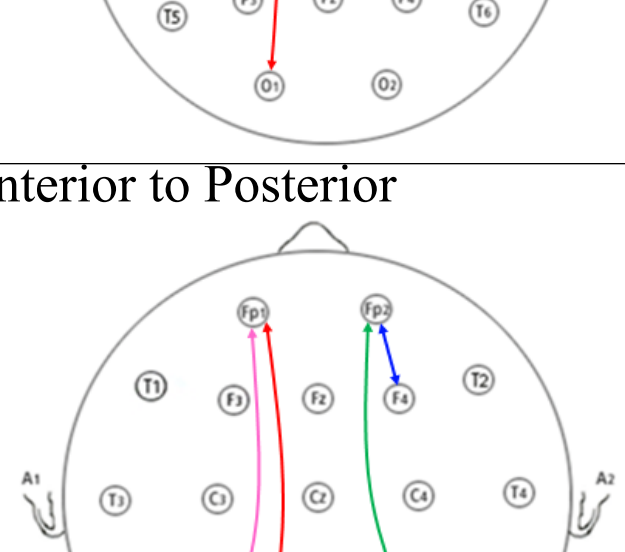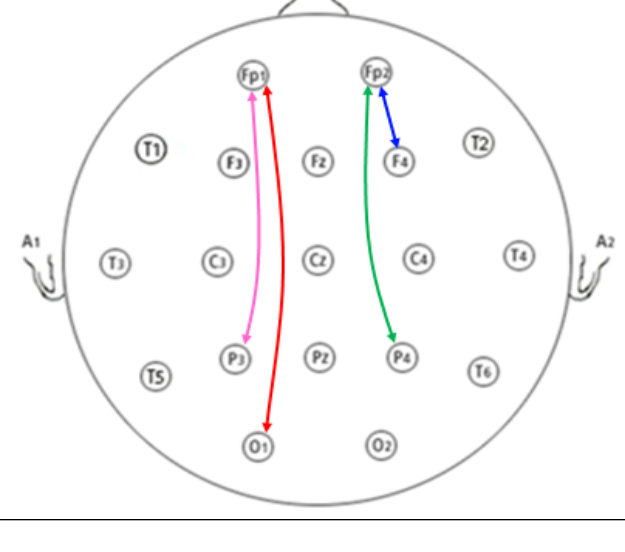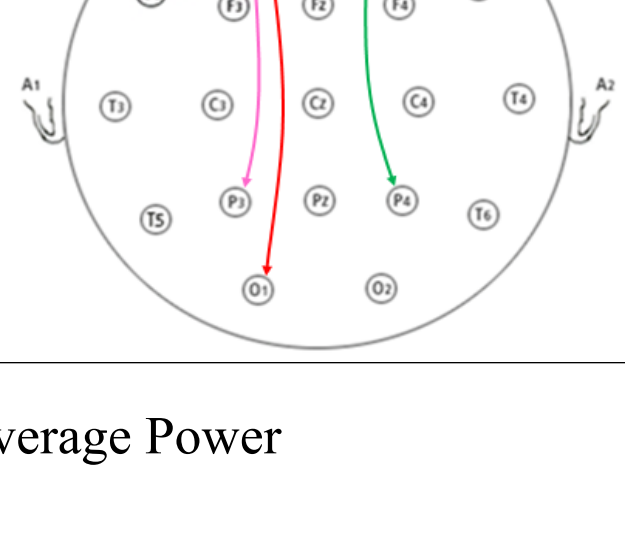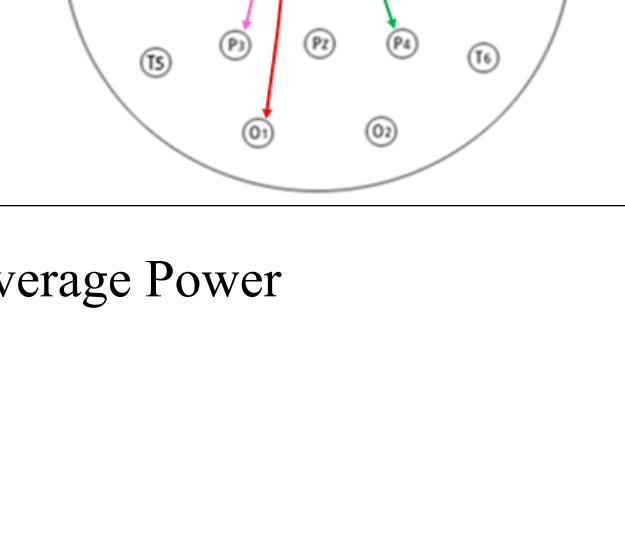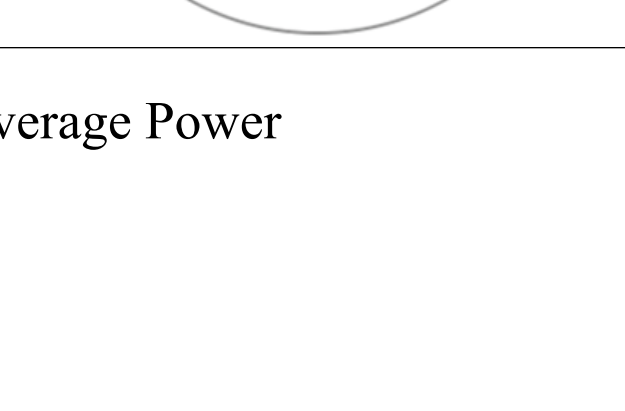
| Channel Pairs | SFFS | t-test |
| --- | --- | --- |
| All | 93.75% | 78.13% |
| Far | 81.25% | 71.88% |
| Anterior | 87.50% | 59.38% |
| Posterior | 87.50% | 81.25% |
| Posterior to Anterior | 81.25% | 56.25% |
| Anterior to Posterior | 68.75% | 43.75% |
| Average Power | 84.38% | 62.50% |



Figure 2– Sets of channel pairs known to have a significant relationship to white matter tracts.

## Discussion

The SFFS algorithm selected features that resulted in classification rates ranging from 69% to 94% overall classification rates, depending on the region that was analyzed for coherence. Table II illustrates the coherence pairs selected by the SFFS algorithm and the corresponding classification rates. The best classification rates (88%) from the known channel pairs were obtained from features taken from anterior coherence, consisting of primarily frontal and fronto-temporal channel pairs, and posterior coherence, in the temporal-parietal area. Of the anterior coherence features selected by the SFFS algorithm, only one passed the t-test: coherence in the gamma band between channel locations T2 and C4 (p-value of 0.02). In the temporal-parietal region, coherence in the delta band between T5 and C3 (p-value of 0.008) passed the t-test. Lower connectivity is common in between AD patients both locally and globally [4].

Each of the three features selected from all channel pairs are in the delta frequency range resulting in a classification rate of 94%. Table I shows that SFFS selected features result in significant improvements in classification accuracy from the classification rates of the four features with the lowest p-value. The Fp1-Fp2 feature selected for all channel pairs has a p-value of 0.086 for all subjects and a p-value of 0.0002 for subjects correctly classified by the feature individually.

Table II – Channel mappings for features selected by SFFS with a maximum of four features for the different channel pairings. The mean and standard deviation of each feature is visible on the right.

| Selected Features | | Controls | | AD | |
| --- | --- | --- | --- | --- | --- |
| | | Mean | SD | Mean | SD |
| **All** | | | | | |
| | Delta C3-T5 | 0.635 | 0.112 | 0.409 | 0.105 |
| | Delta Fp1-Fp2 | 0.682 | 0.137 | 0.729 | 0.134 |
| | Delta F3-T4 | 0.406 | 0.100 | 0.386 | 0.110 |
| **Far** | | | | | |
| | Alpha T6-T2 | 0.243 | 0.033 | 0.210 | 0.045 |
| | Theta T6-F4 | 0.285 | 0.117 | 0.230 | 0.081 |
| | Theta O1-T1 | 0.218 | 0.059 | 0.255 | 0.082 |
| | Theta T5-T1 | 0.226 | 0.062 | 0.220 | 0.061 |
| **Anterior** | | | | | |
| | Alpha T1-C3 | 0.317 | 0.067 | 0.361 | 0.097 |
| | Gamma T2-C4 | 0.273 | 0.088 | 0.368 | 0.127 |
| | Beta Fp1-C3 | 0.461 | 0.113 | 0.402 | 0.099 |
| | Theta F3-C3 | 0.786 | 0.096 | 0.800 | 0.066 |
| **Local Posterior** | | | | | |
| | Delta T5-C3 | 0.612 | 0.112 | 0.409 | 0.105 |
| **Posterior to Anterior** | | | | | |
| | Alpha O1-F3 | 0.279 | 0.073 | 0.254 | 0.073 |
| | Alpha O1-Fp1 | 0.246 | 0.063 | 0.223 | 0.050 |
| | Alpha O1-C3 | 0.419 | 0.112 | 0.318 | 0.111 |
| **Anterior to Posterior** | | | | | |
| | Alpha Fp1-O1 | 0.246 | 0.063 | 0.223 | 0.050 |
| | Delta Fp2-F4 | 0.675 | 0.120 | 0.705 | 0.068 |
| | Theta Fp2-P4 | 0.286 | 0.096 | 0.253 | 0.091 |
| | Gamma Fp1-P3 | 0.276 | 0.099 | 0.283 | 0.095 |
| **Average Power** | Theta P3 | 0.0031 | 0.0001 | 0.0031 | 0.0001 |
| | Delta Fz | 0.0131 | 0.0005 | 0.0132 | 0.0005 |
| | Delta F4 | 0.0016 | 0.0001 | 0.0017 | 0.0001 |
| | Delta O2 | 0.0019 | 0.0002 | 0.0019 | 0.0002 |

## Conclusions

The data driven model that we have employed identifies groups of features within user-determined limits that yield the best classification rates. Based on our results, we can identify networks associated with the frontal-temporal area as being those that exhibit the ability to best separate AD patients from controls when coherence is used as a variable. The networks that were identified as having the highest classification rates as determined by SFFS selected coherence pairs were associated with channels in close proximity – this could imply that the networks involving shorter connections (fibers) are more obviously affected by Alzheimer's disease than those longer frontal-parietal connections, which do not act as variables that are effective in classification.

The SFFS-selected features resulted in higher overall classification rates for the known channel pairs than all of the features using a t-test. This suggests that in a heterogeneous group environment such as a group of patients with varying degrees of Alzheimers, going beyond the t-test with a selection process such as SFFS could prove to be superior.

## References:

[1] Andrew F. Leuchter (1987), 'Electroencephalographic Spectra and Coherence in the Diagnosis of Alzheimer's-Type and Multi-infarct Dementia: A Pilot Study', *Arch Gen Psychiatry*, vol. 44, no.11, pp. 993-998.
[2] A. Delorme and S. Makeig (2004), 'EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics', Journal of Neuroscience Methods, vol. 134, pp. 9-21.
[3] S. Nakariyakul, (2008), 'Improved forward floating selection algorithm for feature subset selection', Wavelet Analysis and Pattern Recognition, vol.2, pp.793-798, 30-31.
[4] Willem de Haan (2009), 'Functional neural network analysis in frontotemporal dementia and Alzheimer's disease using EEG and graph theory', BMC Neuroscience, vol. 10, pp. 101