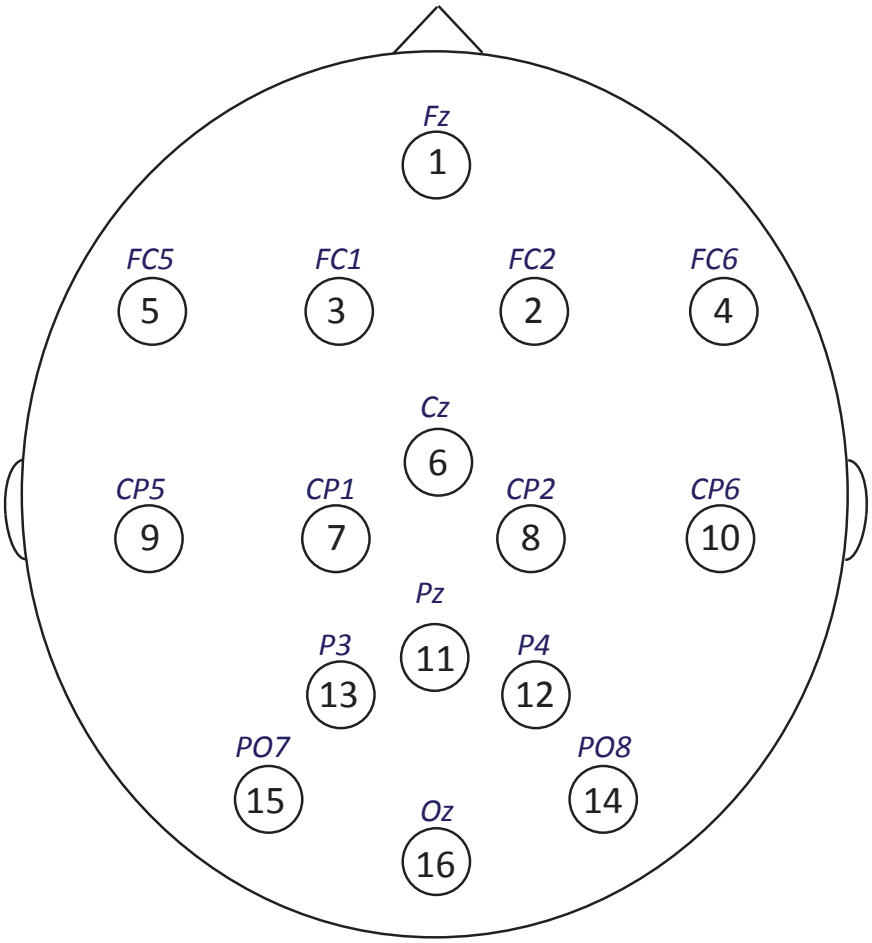
COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK

Lab Instruction 5
P300 Speller

Brain Computer Interface Lab
ECBM 4090



Some materials adapted from g.tec medical engineering (www.gtec.at).

P300 Speller

Our goal in this experiment is to run the P300 Speller. This paradigm is used to help people with movement disorders communicate. Users focus on a target character in a rapidly flashing keyboard. The flashes produce a peak in the Event Related Potential (ERP) approximately 300ms after the onset of the target. We will train a classifier to discriminate between target vs. non-target characters using ERPs, and then test the classifier in a spelling task. In the homework, you will look at the EEG data and calculate the average evoked response (ERP) to the target vs. non-target characters. Ideally, only the target character flashes should produce a P300 waveform.

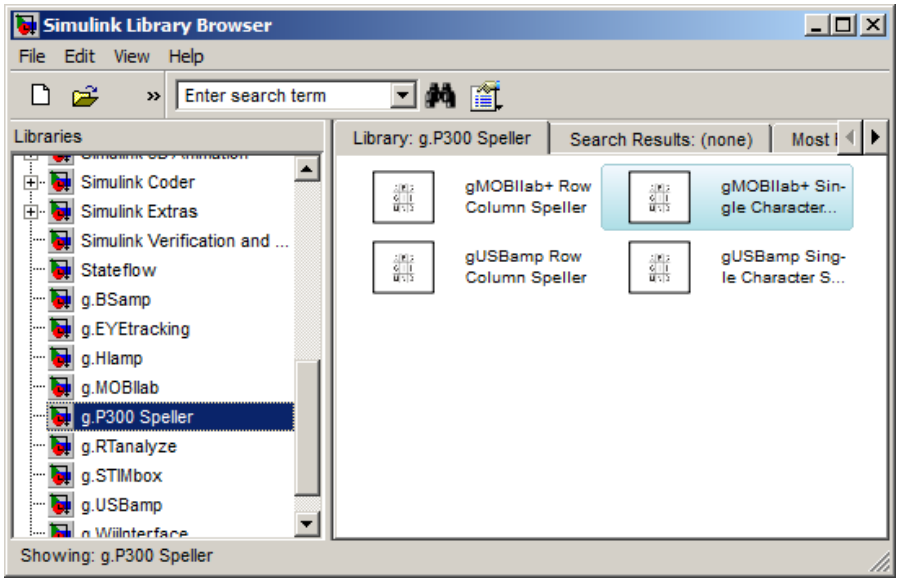
There are three phases in this experiment:

Phase 1. Collect Training Data
Phase 2. Train the Classifier
Phase 3. Test the Accuracy of the BCI System

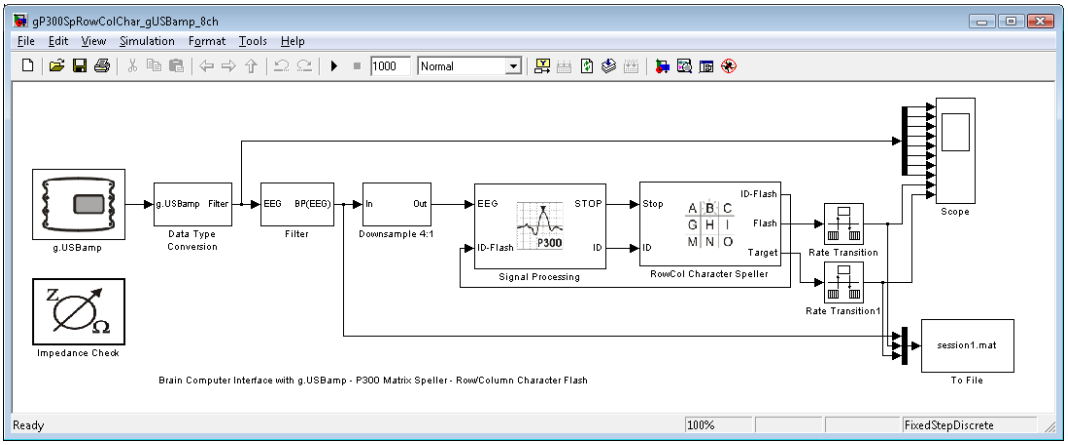
Phase 1. Collect Training Data

Each row and column of letters will flash in a random and sequential order for a specified amount of time. The subject's task is to focus on a specific letter and mentally count how many times this letter flashes.

1. Load the module.
    a. Open the **g.P300 Speller** library listed in the **Simulink Library Browser.**

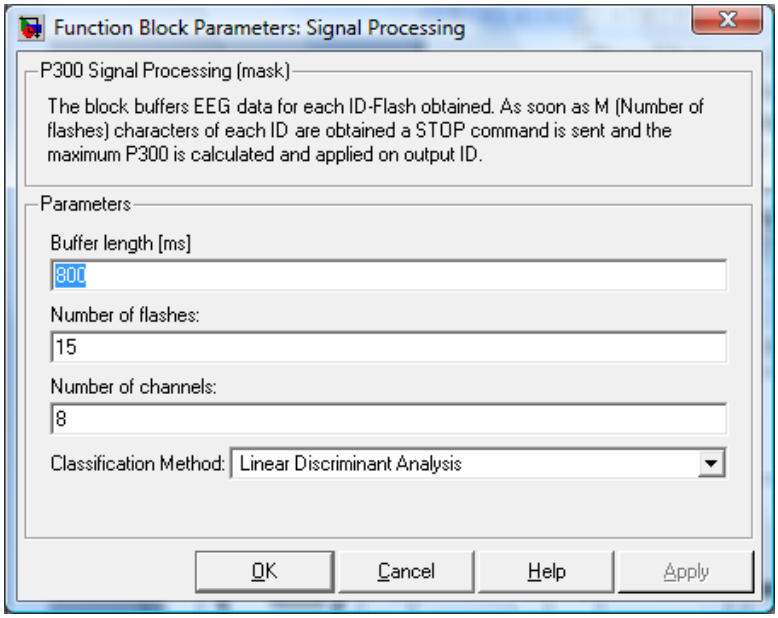


    b. Choose the **g.USBamp Row Column Speller** module. The following Simulink model opens:



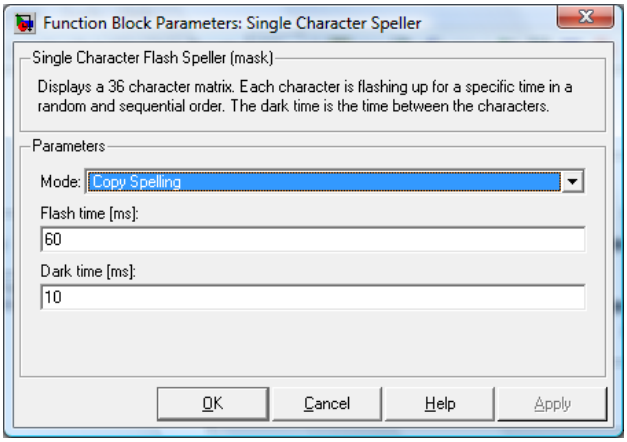
2. Configure g.USBamp.
    a. Double click on the g.USBamp block to open the Configure window.
    b. Select the following electrodes: **Fz, Cz, P3, Pz, P4, PO7, Oz, PO8**.
    c. Select common ground and reference for all groups, set the frame length to 1, and apply a notch filter at 60Hz to all channels.

3. Configure experimental parameters in the module.
    a. Double click on the **Signal Processing** module and set the parameters as follows:



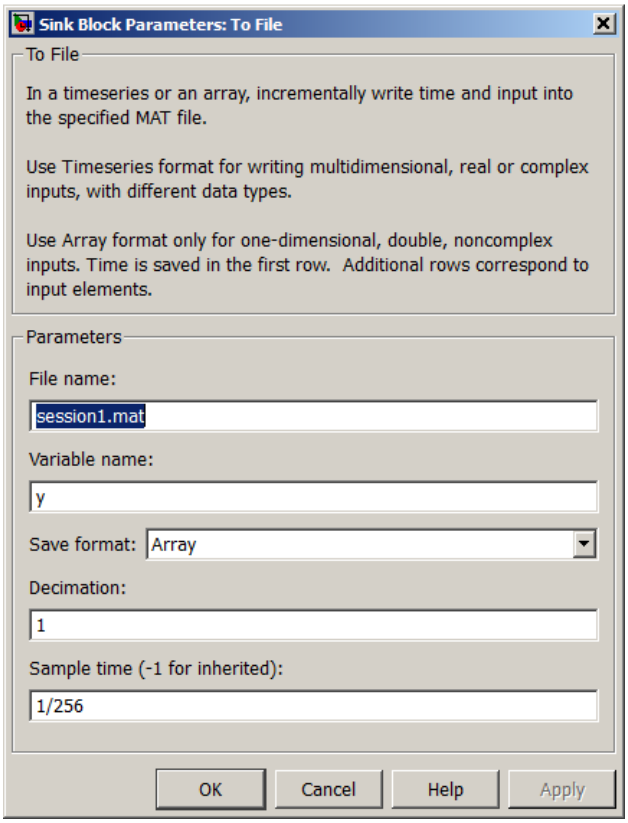
    i. Note: Classification Method does not matter while you are collecting data. You will need to pay attention to this option in the next phase.
    ii. Number of flashes is ‘M’, where each character flashes M times.
        1. M should be set between 10 and 20 depending on your subject.
        2. All M flashes are used to identify the character with the best P300 response, so classifier accuracy will be higher if you select a higher level for M. However, the higher the level of M, the more time it will take to spell a single character.
        3. In general, the goal is to train the classifier for high accuracy with as small an M as possible. In this experiment, we will use an M of 15.

b. The output of the **Signal Processing** block is connected to the **Single Character Speller** or **RowCol Character Speller** block, a MATLAB s-function which controls the experimental paradigm. Double click on it to open the following window:
    i. The window allows you to select between two modes: *Copy Spelling* and *Free Spelling*. Flash time is how many ms the character is highlighted on the screen. Dark time is the time between two flashes in ms.



    i. Select **Copy Spelling** mode.
        a. The labels will be saved along with the data used in the training.
    ii. Set flash time to 100ms, and dark time to 75ms.

4. Collect training data.
    a. Double-click the **To File** block:



i. Enter *train1.mat* under **Filename.**
ii. Enter *y* under **Variable Name**. This stores the data into matrix *y*.
iii. Press **OK** to close the window.

b. **Start** the P300 paradigm in the Simulink module menu.
c. Enter a target word.
   i. The copy spelling mode allows you to enter letters to be copied during the experiment. Enter a word 5-10 characters in length by clicking each character.
   ii. The more letters used for the classifier generation, the higher the accuracy of the feature classification algorithm.
d. Press **Start** to begin the experiment.
   i. The subject should focus on a specific letter while silently counting that letter's flashes
   ii. Close the window when all characters have been recorded by the subject.
   iii. When you are collecting training data, the program may predict incorrect letters. This is to be expected since you have not yet trained the classifier.

Phase 2. Training the Classifier

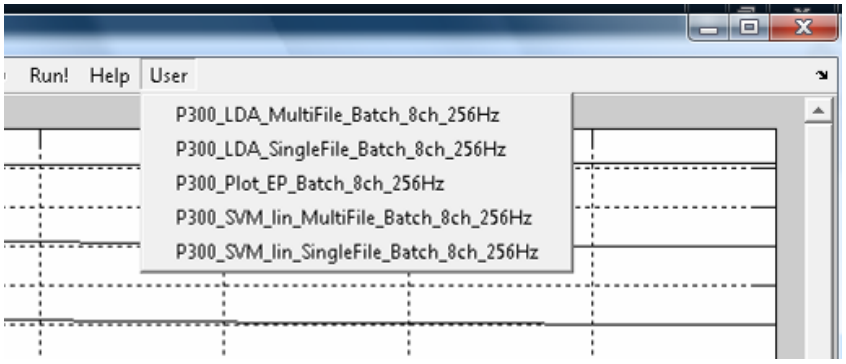
The P300 Speller uses *Linear Discriminant Analysis* or *Support Vector Machines* for feature classification. To use these signal processing algorithms, you first have to generate classifiers.

1. To start the Data Editor, type *gbsanalyze* in the MATLAB command window.

2. Load the acquired data *train1.mat* from the copy spelling. The sampling frequency is 256 Hz.

3. Select **Appearance Settings** from the Options menu and set **User Directory** to C:/Program Files/gtec/gtec_libraries/g.p300/batch.

4. Select **P300_SVM_lin_SingleFile_Batch_8ch_256** from the User menu. The classifier is saved to file and loaded automatically to the MATLAB workspace for all batches.

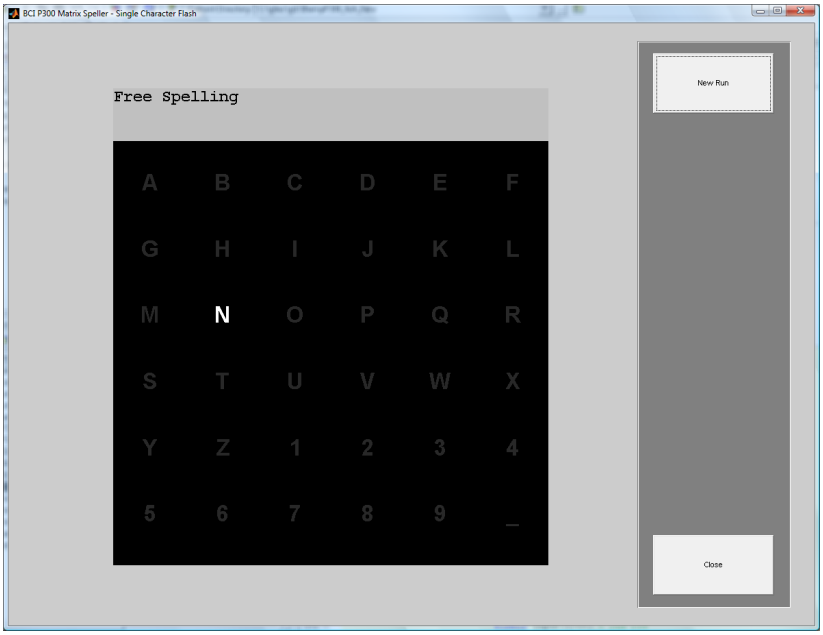


5. This data (train1.mat) is required for the homework and MiniProject. Make sure to change the filename in To File before running the next session to avoid overwriting data!

Phase 3. Testing the Accuracy of the BCI System with Online Classification

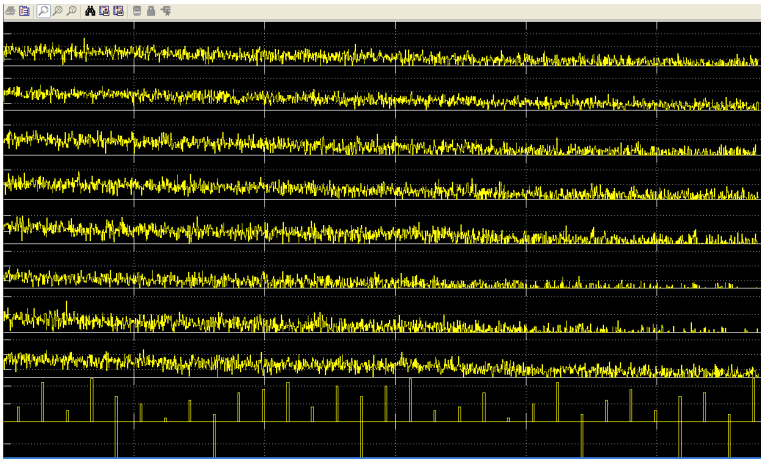
1. Configure the Simulink module for testing.
    a. In the **Signal Processing** block, select the classification method you want to use. The classifier must be present in the workspace, and should have been loaded by g.BSanalyze.
    b. Remain in **Copy Spelling** mode in the Speller block.
    c. In the **To File** block, change the filename to *test1.mat*.
        i. This data will be used in the MiniProject

2. Run the testing phase.
    a. Start the Simulink module.
    b. The window will flash each character in random order. The subject must concentrate on the specific character he or she wants to spell. After the Signal Processing module determines the character, the character is shown on the paradigm window, allowing the subject to write words without any previous input.



Report: Spell the name of your group (e.g., Group_2) and take a screenshot. (10 pts)

3. Familiarize yourself with data storage and visualization (to assist with HW and MiniProject)
    a. The **Speller** block output signals are called **Flash** and **Target**.
        i. **Flash** is the ID of the character that has flashed.
        ii. **Target** is a trigger signal indicating when the target letter has flashed in **Copy Spelling** mode.
        iii. **To File** stores the EEG data from **Flash,** and **Target**.
    b. Double-click onto the Scope block to investigate these signals:
        i. Channels 1-8 are the EEG data.
        ii. The 9th channel contains the ID of the character that has flashed. It shows the integers 1 to 36 of the 36 characters that can be displayed for Single Character Speller block and 1-12 of the different rows/columns for the Row/Col Character Speller block.
        iii. The 10th channel indicates the time point when the target letter has flashed. It displays 1 if the target character has flashed.



Homework: Display and compare the ERP for target vs. non-target flashes.

Channel ten in the stored variable shows when a signal has flashed and which row and column were flashed. Because the keyboard used is 6 x 6, as seen above, values 1-6 in channel ten identify the row, and values 7-12 identify the columns (so 4 → row 4 and 10 → column 4). For example, if the target character is 'N', it means when channel ten is either 3 (row) or 8 (column), the target character 'N' has flashed.
Each time the row or column of the target character flashes, it is stored in channel eleven.
Combine this information to segment the *train1.mat* EEG data into target and non-target flashes.

• Use channels ten and eleven as triggers to cut the EEG data. (1 pt)

• Cut 100ms before onset through 600ms after each flash and separate target and non-target characters. (2 pts)

• Take the average response over trials and display the ERP waveforms for a) target and b) non-target conditions for each channel and the average signal over all channels. (6 pts)

• Do you see a P300 response to the target? (1 pt)

Hint: To find the onset of the flashes, you can use the following MATLAB script:

```
Index = find (diff ( trig) > 0 );
```

* Channel 11 contains target information only in Copy Spelling, not Free Spelling.