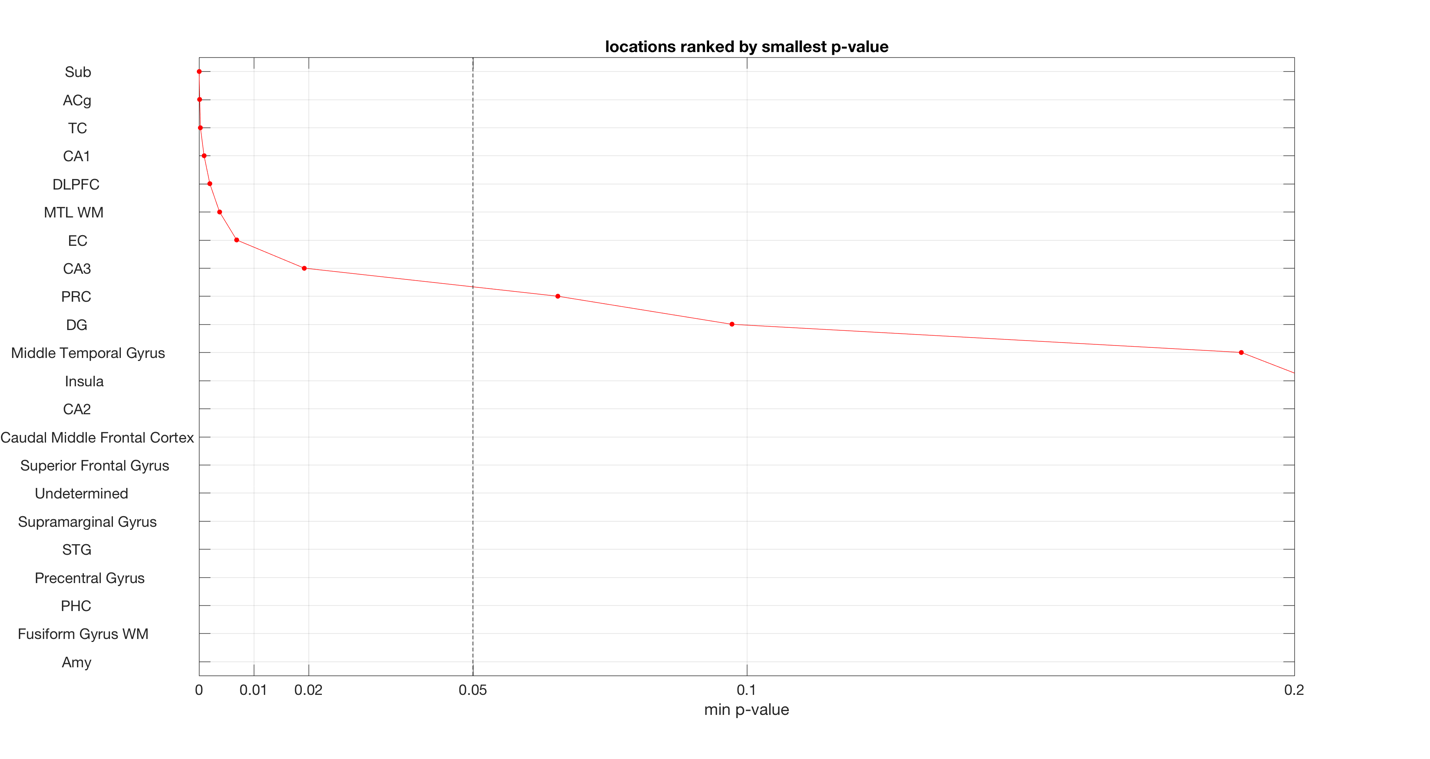
# Methods

Consider sham stimulation to be your baseline mean/stderr (any sample with amplitude=0). For every stimulation sample tuple (location, frequency, amplitude, duration), it is interesting if the distribution of biomarker changes is significantly different from the sham distribution. This hypothesis is tested via two-sided t-test, ie. do the sham samples and stimulation samples come from the same distribution?

We must correct for multiple comparisons at each location. For all the samples drawn from a location, use the Bonferroni-Holm correction which adjust p-values so that instead of a 0.05 rate of false positives, we get a false-detection rate of 0.05. This is a common approach and is less severe (conservative) as the original Bonferroni approach of dividing 0.05 by m, where m is the number of repeated samples. From here on, we refer to corrected p-values.

## Rank by minimum p-value

For each location, assign that location the minimum p-value (corrected, two-sided) found among all samplings at that location.



We can see that the “Sub” location had the most unusual observation. Let’s look more closely at the smallest three p-values for that location:

Location Frequency Amplitude Duration Subjects Samples Pvalue

\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_

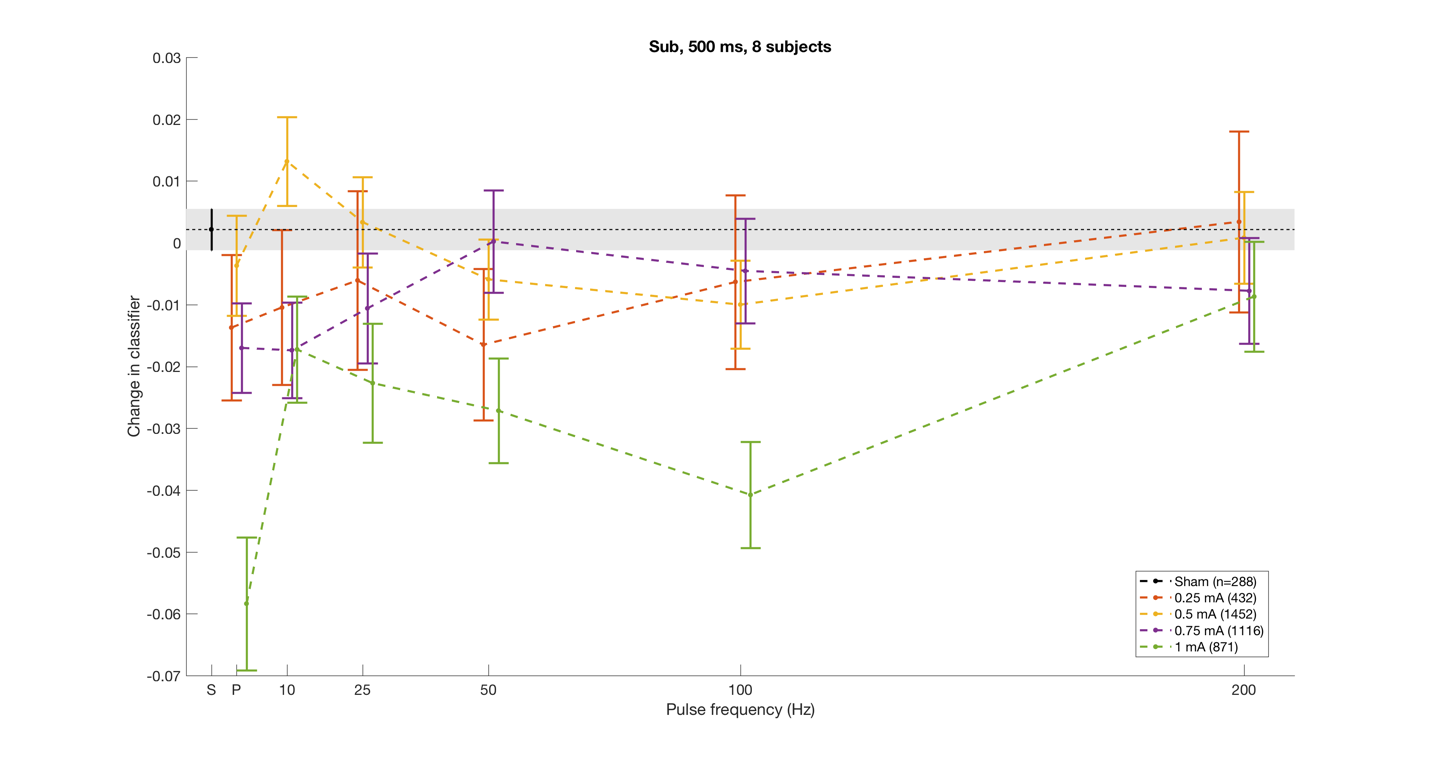
Sub 0 1 0 4 120 1.2263e-07

Sub 100 1 500 4 150 1.6384e-06

Sub 50 1 500 4 150 0.00069874

So the smallest value that it was ranked by was a pulse at amplitude=1

Let’s look at that plot:



The set of points of interest are in the lower left (light green error bars). In other words, those are the most significant outlying points among all locations (minimum p-value of all p-values).

Let’s look at the second: Acg

Location Frequency Amplitude Duration Subjects Samples Pvalue

\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_

ACg 200 0.75 500 1 156 2.4612e-06

ACg 50 1 500 3 192 0.00052003

ACg 200 1.5 1000 1 30 0.0027358

ACg 100 0.75 500 1 156 0.0031033

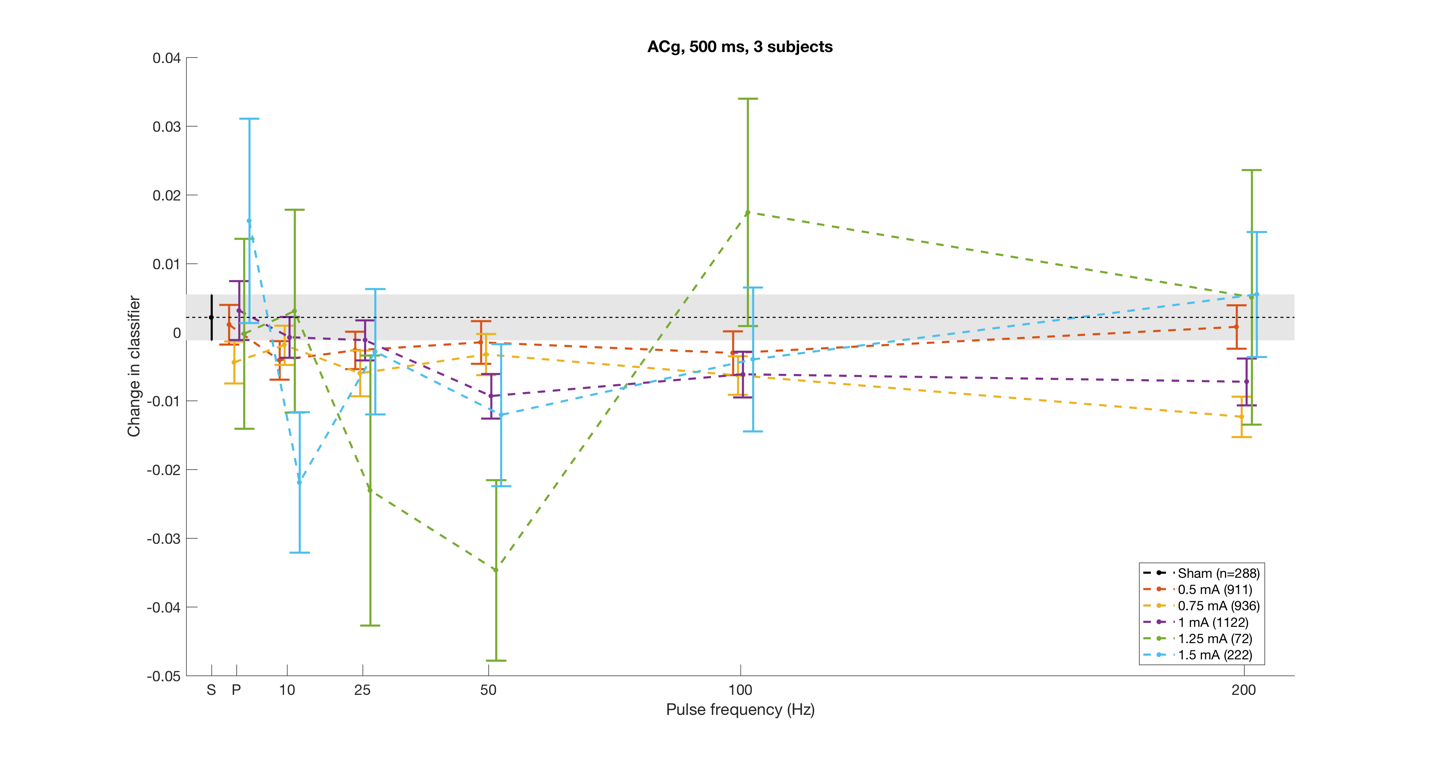
ACg 200 1 500 3 192 0.006739

ACg 100 1 500 3 192 0.013589

ACg 50 1.5 250 1 30 0.014925

ACg 25 0.75 500 1 156 0.01551

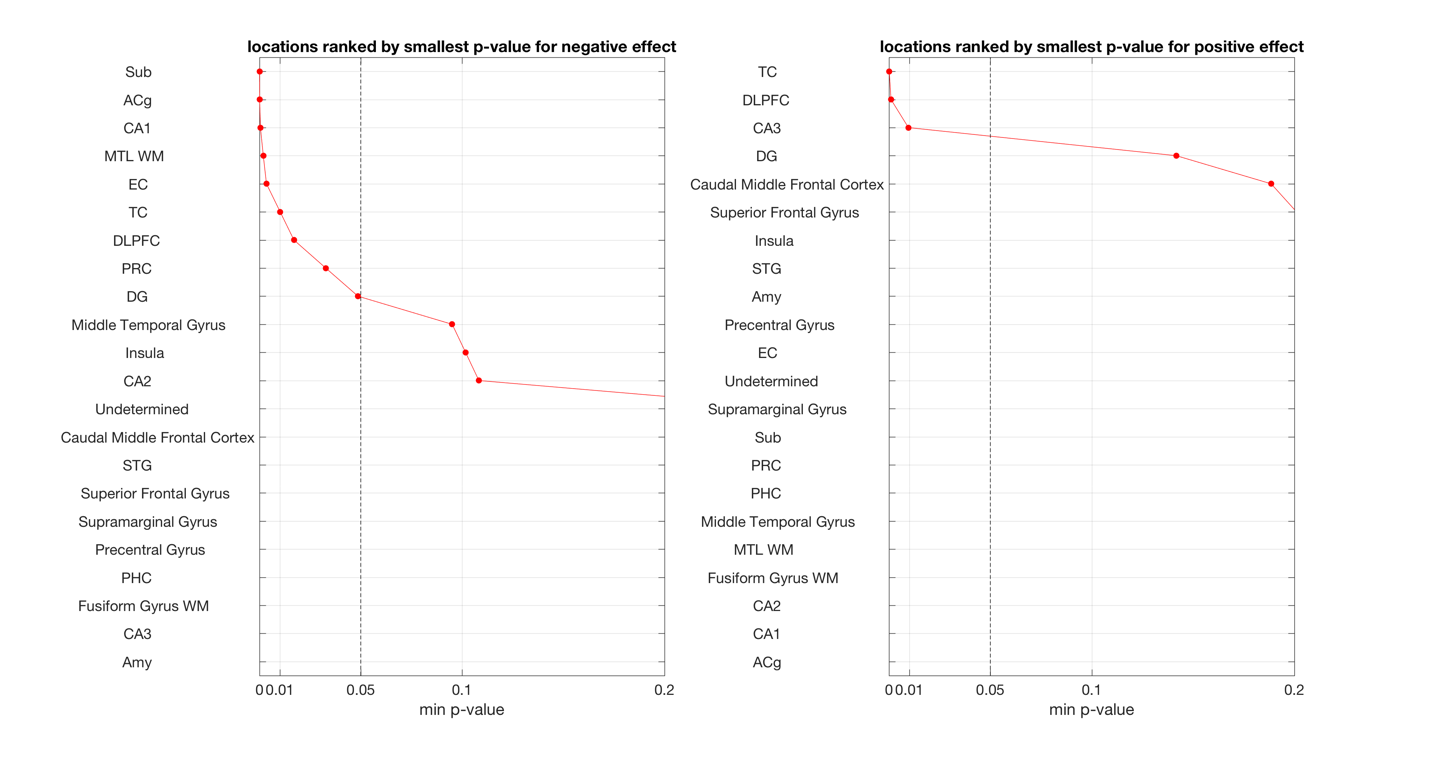
ACg 50 1.25 500 1 12 0.017155

ACg 10 1.5 500 1 42 0.023948

The most interesting stimulation was 200 Hz, 0.75 mA (yellow). It has a tight standard error. What’s odd is that we maybe would have expected the 50 Hz, 1.25 mA (green) to have been more interesting, but because it has such a large standard error, it is less confident in that result.

## Stratify by positive/negative effect, rank by minimum p-value

The goal is to look separately at the samples with the most significant positive effect and samples with the most significant negative effect. Instead of using the two-sided t-test as above, we use the one-sided t-test (right for positive, left for negative). We expect to see the above examples in the table of most significant negative effects.



Some observations:

* More areas are negatively affected than positively affected
* CA1 is negative while CA3 is positive

We’ve already examined “Sub” and “ACg” to confirm their negative effect, so now let’s look at TC to examine the positive effect:

Location Frequency Amplitude Duration Subjects Samples Pvalue

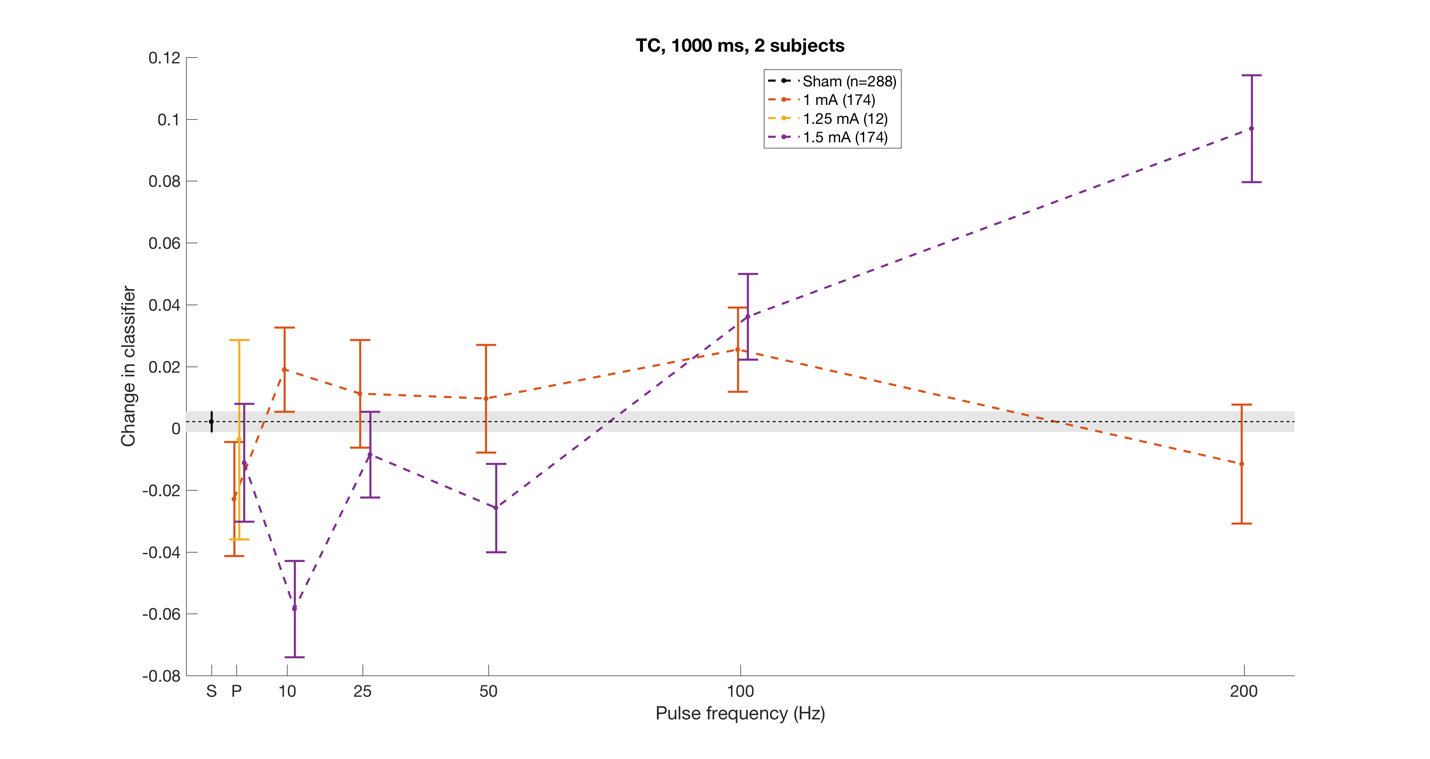
\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_

TC 200 1.5 1000 1 30 3.2861e-06

TC 200 1.5 500 1 54 0.00016388

TC 100 1.5 1000 1 30 0.010291

The first and third are both duration=1000. Let’s look at them together:



This suggests that it might be interesting to explore 150 Hz, for example.

## Ranking an entire location based on all of samplings at that location

For each tuple of (loc,freq,amp,dur), the p-value is the probability of the biomarker changes being “interesting”: P(L,f,a,d). The probability of a whole location being interesting is the joint probability of seeing the entire set of observations at that location: prod[P(f,a,d|L)] for all combinations of (f,a,d). In other words, the product of all the p-value calculations at a location. Again use one-sided t-test since ranking positive and negative separately.

